The Trolley and the Sorites

John Martin Fischer

I

In two fascinating and provocative papers, Judith Jarvis Thomson discusses the "Trolley Problem."[1] In this paper I shall present my own version of the problem. I explain the problem and its apparent significance. Then I present a series of cases which calls into question the intuitive judgments which generate the original problem. Thus, I propose to *dissolve* the problem. The method of argumentation that issues in this dissolution might be considered troublesome. For this reason I explore certain objections to the methodology. Finally, I undertake to explain the significance of the dissolution of the problem.

II

Let us follow Thomson in calling the first case "Bystander-at-the-Switch." A trolley is hurtling down the tracks. There are five "innocent" persons on the track ahead of the trolley, and they will all be killed if the trolley continues going straight ahead.[2] There is a spur of track leading off to the right. The brakes of the trolley have failed, and you are strolling by the track. You see that you could throw a switch that would cause the trolley to go onto the right spur. Unfortunately, there is one

1. "Killing, Letting Die, and the Trolley Problem," *The Monist* (April 1976): 204-17; and "The Trolley Problem," *The Yale Law Journal* 94 (May 1985): 1395-1415. Both articles are reprinted in William Parent, ed., *Rights, Restitution, and Risk* (Cambridge: Harvard University Press, 1986), 78-116. Thomson attributes the original formulation and discussion of the problem to Philippa Foot, "The Problem of Abortion and the Doctrine of the Double Effect," *Oxford Review* 5 (1967): 5. I have discussed this set of issues in John Martin Fischer, "Thoughts on the Trolley Problem," in John Martin Fischer and Mark Ravizza, eds., *Ethics: Problems and Principles* (Fort Worth, Texas: Harcourt Brace Jovanovich, forthcoming 1991); and "Tooley and the Trolley," *Philosophical Studies* 62 (1991): 93-100; and in John Martin Fischer and Mark Ravizza, "Thomson and the Trolley" (manuscript); and "Quinn on Doing and Allowing," *Philosophical Review* (forthcoming). There is also a selection of articles pertaining to the Trolley Problem in Fischer and Ravizza, eds., *Ethics*.

2. I shall use "innocent" in a rather broad sense to mean that the persons have not done anything morally wrong for which they "deserve" to die or have forfeited their right to life, etc. Further, I shall assume throughout this paper that no "special" facts distinguish the persons involved in the cases. For instance, none of the six persons involved in "Bystander-at-the-Switch" is a mass-murderer, great scientist, and so forth; further, one has not made special arrangements with any of them. Of course, I shall be engaging in schematic and partial descriptions of the various cases, and the reader will need to keep in mind that certain ways of filling in the details of the cases will affect the moral judgments appropriate to the cases.

innocent person on that spur of track. You are an "innocent bystander," *i.e.*, not an employee of the railroad, and so forth. You can throw the switch, saving the five persons but causing the one to die, or you can do nothing, thereby allowing the five to die. What should you do?

Well, it seems that it would at least be permissible for you to turn the trolley to the right, thus saving the five but killing the one. Perhaps it is also obligatory to do this, but it is at least intuitively plausible that one *may* turn the trolley to the right.

But consider now a second case, "Fat Man." You are standing on a bridge watching a trolley hurtling down the tracks toward five innocent persons. The brakes have failed, and the only way in which you can stop the train is by impeding its progress by throwing a heavy object in its path. There is a fat man standing on the bridge next to you, and you could push him over the railing and onto the tracks below. If you do so, the fat man will die but the five will be saved.

What ought you to do? Whatever else is true about the situation, it seems that it would be impermissible for you to push the fat man over the railing. Thomson emphasizes this point, saying, "Everybody to whom I have put this case says it would not be [permissible]."[3] But now the question arises, why is it permissible to save the five in "Bystander" but not in "Fat Man"? By reference to what general principle (or principles) can the two cases be distinguished morally?

Consider a third case, "Transplant." Imagine that you are a surgeon—a truly great surgeon. Now there are five persons in the hospital, each of whom needs an organ in order to survive. It just happens that an innocent visitor has arrived in the hospital, and you know that he is tissue-compatible with all the people who need organs, and that you could cut him up and distribute his parts among the five who need them. Would it be permissible for you to perform the operation (without his consent)? It seems quite evident that it would not be permissible for you to proceed. But why is it permissible for you to save the five in "Bystander" but not in "Transplant"? Why exactly are "Fat Man" and "Transplant" morally similar to each other but crucially different from "Bystander"? This is, I believe, a particularly perspicuous formulation of the Trolley Problem.

The Trolley Problem can be seen to be a problem for both major kinds of ethical theory: consequentialism and deontology. Consequentialism enjoins the maximization of some impersonally defined good.[4] A certain sort of consequentialist would seem to have a difficult time explaining why one could save the five in "Bystander" but not in "Fat Man" and "Transplant." For it seems that a consequentialist would need to pre-

3. Thomson, "The Trolley Problem", in Parent, ed., *Rights*, 109.

4. For a useful definition and analysis of consequentialism, see Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Clarendon Press, 1982).

scribe saving the five in *all* the cases. The Trolley Problem is also a problem for the deontologist, who believes there are constraints on the maximization of the good. A certain sort of deontologist will want to agree with the intuitive judgments presented above according to which one may save the five in "Bystander" but not in "Fat Man" and "Transplant." Such a theorist will need to generate a principle which would have these results. Another sort of deontologist might wish to deny that it is permissible to save the five in *any* of the cases. But even this sort of theorist will need to explain away the evident plausibility of the intuition that one *may* save the five in "Bystander."

III

Upon first considering the above cases, I agreed with the pattern of intuitive judgment which produces the Trolley Problem. I took the challenge to be to generate a (deontological) principle or set of principles which would explain and suitably systematize the constellation of intuitions. But having reflected on the examples, a rather startling thing happened: I realized that a set of considerations which seem quite unassailable might lead one to think that all of the cases are morally on a par. That is, despite the apparent moral differences between the cases presented above (and similar cases), I began to think that there was a way to establish that the cases were morally similar at a deep level. The method of argumentation for this admittedly surprising—and alarming—assimilation of seemingly disparate cases proceeds via the construction of intermediary examples.

Before presenting this set of considerations, I pause to explain what I mean by the crucial phrases, "morally similar," "morally equivalent," "morally on a par," and so forth. In this paper I am using these phrases in a somewhat "technical" sense. That is, by these phrases I do *not* mean to indicate total moral sameness. Rather, I mean to indicate that *certain* key moral claims are the same in the cases. More specifically, I mean to indicate that the claims about the permissibility or impermissibility of saving the five are the same in each case. Thus, in this paper, when I claim that two cases are appropriately assimilated morally, I shall not mean that they are properly treated as the same in *all* moral respects; rather, I shall mean that the claims about the permissibility or impermissibility of saving the five have the same truth value. (Put in terms that are compatible with a non-cognitivist approach to ethical language, to say that two cases are to be assimilated morally is to say that it is permissible to save the five in one case if and only if it is permissible to save the five in the other.)

To proceed, let us consider a fourth case, "Ramp." "Ramp" is in many respects quite similar to "Fat Man": there is a trolley hurtling down a track toward five innocent persons, and an innocent fat man

standing on a bridge above the tracks. Let us suppose that the bridge is a railroad bridge (*i.e.*, with railroad tracks on it). And let us imagine that you could push a button that would cause a ramp to go up underneath the trolley. Further, if you were to push the button and thus cause the ramp to go up, you would thereby cause the trolley to jump up to the tracks on the bridge and to continue along those tracks. If you were to push the button, you would save the five, but, regrettably, the trolley would run over the fat man. In "Fat Man" the only way in which you could save the five would be by shoving the fat man onto the trolley. In "Ramp" the only way in which you could save the five would be by "shoving the trolley onto the fat man." In all other respects the cases are the same, and I have a strong inclination to say that "Fat Man" and "Ramp" are morally on a par.

To strengthen my point, consider "Fat Man*." In "Fat Man*" the situation is exactly as it is in "Fat Man," except that you can push a button which would cause the bridge's handrail to wobble, which in turn would cause the fat man to topple in front of the trolley. Although there is a factual difference, I think that there is no moral difference between "Fat Man" and "Fat Man*." And it seems to me quite evident that there is no moral difference between "Fat Man*" and "Ramp." I do not find it normatively plausible to suppose that the factual difference between these two cases underwrites a moral difference between them. How can it make a difference whether one pushes the button that causes the fat man to fall into the path of the trolley or the button that causes the trolley to "jump" toward the fat man? Thus, because there is no reason to think that the pertinent judgments are not transitive, it should be evident that there is no moral difference between "Fat Man" and "Ramp."

Perhaps the reader will not have failed to notice a rather disconcerting implication of the above remarks. I claim that "Fat Man*" and "Ramp" are morally on a par, and thus that "Fat Man" and "Ramp" are as well. But "Ramp" is clearly morally on a par with "Bystander." Thus it would seem that I am committed to saying that "Bystander" is morally on a par with "Fat Man," and this result is rather alarming indeed. After all, it is a presupposition of the Trolley Problem that it would be permissible to save the five in "Bystander" but not in "Fat Man."

I am willing to accept this result because there is a plausibility argument for it that is implicit in what I have said above, and I do not see how to challenge this argument. There are two steps, each of which seems unassailable. First, "Bystander" is morally similar to "Ramp," because whether the trolley is shunted to the right or upward cannot be morally significant. Second, "Ramp" is morally similar to "Fat Man," because whether one throws the trolley onto the fat man or throws the

fat man onto the trolley cannot be morally significant. Thus, it is hard to deny that "Bystander" is morally on a par with "Fat Man."

There is another route to the conclusion that the two cases must be assimilated. Consider the following case, which I shall call "Jiggle-the-Tracks I." A trolley is hurtling down a track toward five innocent persons. As before, the brakes have failed, and there is no way of stopping the trolley. You are a bystander watching the situation develop. You notice that there is a spur of track leading to the right on which (very close to the bifurcation of the tracks) stands one innocent person. You also notice that you could push a button which would cause the tracks to "jiggle" in such a way as to cause the trolley to go onto the right fork. If you were to push the button, you would save the five persons but cause the death of the one.

It seems to me that "Jiggle I" is morally on a par with "Bystander." That is to say, if it is permissible to save the five in "Bystander," surely it is permissible to save the five in "Jiggle I": there is no moral difference between redirecting the trolley via the switching mechanism envisaged in "Bystander" and the jiggling envisaged in "Jiggle I."

Imagine now a related case, "Jiggle-the-Tracks II." Everything is the same here except for the manner of track-jiggling that would result from your pushing the button. That is, there is a trolley hurtling down the tracks toward five persons, with a spur leading to the right on which there stands a man (very close to the point of track bifurcation). But this time if you push the button, the tracks would be jiggled in such a way that the man would be toppled onto the track in front of the trolley. We can suppose that the man is large enough to stop the progress of the trolley, but that if this were to occur, he would be killed.

It seems to me that this case is morally parallel to "Jiggle I" in the sense that it would be morally permissible to push the button in "Jiggle I" if and only if it would be morally permissible to do so in "Jiggle II." How can it make a moral difference whether one jiggles the trolley onto the man or the man onto the trolley?

Further, it should be apparent that "Jiggle II" is morally on a par with "Fat Man*." In "Fat Man*" the man is jiggled *down* onto the train. In "Jiggle II" the man is jiggled *sideways*. But surely *this* factual difference cannot make a moral difference. As was argued above, "Fat Man*" is morally on a par with "Fat Man." By transitivity, then, we have arrived (again) at the assimilation of "Bystander" and "Fat Man."

By reasoning parallel to that which implies that "Fat Man" is morally equivalent to "Bystander," one can show that "Fat Man" is morally equivalent to "Transplant." That is, if one believes that "Fat Man" is morally equivalent to "Bystander," one should also believe that "Fat Man" is morally equivalent to "Transplant." To see this, begin by think-

ing about the following example, "Shield."[5] Suppose that everything is like the situation in "Fat Man" except for the following. You are a great surgeon and you are standing next to the (unfortunate) fat man. You cannot lift him, and he is so positioned that you cannot cause him to fall in front of the trolley. But it turns out that you can (painlessly—your technicians and nurses are with you) remove the fat man's inner organs and construct a shield with them that would stop the trolley. Imagine that this is the only way in which you could stop the train, and that the fat man would die in the process.[6]

It seems to me that "Shield" is morally on a par with "Fat Man." That is, if it is permissible to save the five in "Fat Man," it is also permissible to save the five in "Shield." How could the difference between throwing the fat man in front of the trolley and constructing a shield out of his organs make a *moral* difference?

But "Shield" is clearly morally on a par with "Transplant" because it involves taking one person's organs to save five people. So, it seems that we must conclude that "Fat Man" is morally on a par with "Transplant." Since we have already seen that "Bystander" and "Fat Man" are morally on a par, we can conclude via transitivity that "Bystander" and "Transplant" are morally on a par. That is to say, if it is morally permissible to save the five in "Bystander," it is morally permissible to save the five in "Transplant." Thus, we have reached the result that all three cases—"Bystander," "Fat Man," and "Transplant"—ought to be assimilated morally, despite one's first intuitive reaction that they should not be assimilated with regard to the relevant permissibility claims.

The assimilation of "Bystander" and "Transplant" has proceeded via the intermediate case, "Fat Man." But an argument for this assimilation need not employ "Fat Man." Consider the following intermediate case, "Scan."[7] Again, as in "Transplant," there are five people dying—this time they are dying of some disease. Let us suppose that there is a scanning device that can scan the brain of an individual (with certain neurological properties) and generate information that can be used to save the five dying persons. Imagine further that there is an innocent visitor to the hospital who is neurologically suited to the scan. Unfortunately, if you do the scan, you kill this innocent person, for the scanning method is lethal. You can save the five if you do the scan, but you will kill the one if you do so, and this is the only way in which you could save the five.

It seems to me that "Bystander" and "Scan" are morally similar (as regards the permissibility claim). In "Bystander" you can save the five

5. This example was presented to me by Gary Watson.

6. Ernest Partridge has pointed out to me, apropos of this example, that it can take guts to stop a train. Also, both Dave Estlund and Greg Ravizza have suggested a variant of "Shield," "Lasso," in which you make a lasso out of the poor fat man's organs and thus stop the train.

7. "Scan" is presented and discussed in Fischer, "Tooley and the Trolley."

only by directing the train onto the one; in "Scan" you can save the five only by directing the lethal rays onto the one. Is there a moral difference between a lethal trolley and lethal rays? It does not seem that there is any factual difference that can plausibly be thought to make a *moral* difference between these two cases.[8] Further, "Scan" seems to be morally similar to "Transplant." If these two claims are true and transitivity holds for the relevant notion of moral similarity, then there is reason to hold that "Bystander" and "Transplant" are morally similar. The construction of an intermediate case, "Scan," thus leads to the moral assimilation of "Bystander" and "Transplant," quite independent of considerations pertinent to "Fat Man."

Imagine, also, the following case, which I shall call "Cart." There are five healthy and innocent people in a waiting room in a hospital, and one healthy and innocent person in a waiting room adjacent to the first. There has been an earthquake, and a surgical cart—filled with very sharp surgical instruments—is hurtling down the hall toward the room in which the five people are waiting. Regrettably, you cannot stop the cart, nor can you prevent it from entering the room with the five people, except by deflecting it into the room in which the one person is waiting. Further, let us suppose that the instruments are arranged in such a way that, given the trajectory of the cart and the locations of the people, they would all be impaled if you allowed the cart to proceed. Unfortunately, if you shunted the cart into the other room, the instruments would impale the one person. You cannot save the five without killing the one.

Now it seems to me that there is no moral difference (in the relevant respects) between "Cart" and "Transplant." Although there plainly are many factual differences, I do not see how they could underwrite a claim that it would be permissible to save the five in one case ("Cart") but not in the other ("Transplant").[9] Further, there is obviously no moral difference (in the relevant respects) between "Cart" and "Bystander"—if it is permissible to save the five in "Bystander," it is permissible to save the five in "Cart." Thus, again, we are lead to the conclusion that "Bystander" and "Transplant" are morally on a par.

The method of intermediation has been employed to generate the following results. "Bystander" was alleged to be morally similar to "Fat Man." "Fat Man" was alleged to be morally similar to "Transplant." Thus, it was concluded that "Bystander" is morally similar to "Transplant." Further, independent arguments (not proceeding via "Fat Man")

8. I cannot here exhaustively canvass potential moral principles that would imply a difference between "Bystander" and "Scan" and argue for their ultimate inadequacy. There are, however, arguments against these alternative principles in Fischer, "Thoughts on the Trolley Problem" and "Tooley and the Trolley"; and in Fischer and Ravizza, "Introduction: Moral Problems and Principles" in Fischer and Ravizza, eds., *Ethics*; "Thomson and the Trolley"; and "Quinn on Doing and Allowing."

9. See note 8.

were adduced that show that "Bystander" and "Transplant" are morally similar. The method of intermediation appears to show, then, that all three cases are, at a deep level of analysis, morally similar with regard to the relevant permissibility claims. This method of argumentation, then, challenges the presuppositions of the Trolley Problem. Indeed, it threatens to dissolve the problem.

IV

A.

The method of argumentation employed above—the method of intermediation—generates at least some reasons to assimilate the original three cases which constitute the Trolley Problem. The examples seem to me quite striking and compelling. But the result to which they lead is very surprising and challenging. It is, to put it mildly, highly unintuitive that "Bystander" should be considered morally on a par with "Transplant" as regards the permissibility of saving the five—this is a radical and alarming conclusion. One should therefore be careful about jumping hastily to any conclusion, and one should reflectively scrutinize the methodology. It might be best to look at the argumentation in the above section as posing a challenge. The challenge is to explain away the putative necessity of assimilation by pointing to what is wrong with the argumentation or to reassess one's views about the examples. In the latter case, one ought also to think about the significance of this reassessment; specifically, one ought to attempt to fit one's new views about the examples with one's ethical theory or views about the nature of moral choice.

First, let me say a few words about the use of examples—actual and hypothetical—in thinking about morality. Some philosophers have objected to the use of examples—especially hypothetical examples—in testing ethical theories or in seeking to understand practical reasoning. They have raised various objections to the use of examples in moral theorizing and in seeking guidance in practical affairs. Elsewhere, my coauthor and I have laid out some of these objections and have provided some considerations in defense of the use of examples.[10] Here, I shall not consider these general matters further, except to make the following point. My criticism of the presuppositions of the Trolley Problem is an *internal* criticism in the sense that the Trolley Problem presupposes the relevance of hypothetical examples to ethical theorizing. I am most decidedly not foisting upon the proponent of the Trolley Problem the consideration of certain sorts of hypothetical examples; such a person is already firmly in the tradition which considers such examples relevant.

Now someone might say that the examples I have adduced constitute a

10. See Fischer and Ravizza, "Introduction: Moral Problems and Principles."

reductio ad absurdum of the whole methodology. That is to say, one might think that anyone who is willing to concede the relevance of hypothetical examples to ethics *must* admit the cogency of the argumentation presented above (which appears to demand an assimilation of the puzzle cases). And one might conclude from this that one ought not concede the relevance of such examples to ethics in the first place. As far as I can see, this is an open option, and I do not propose to address it here. Indeed, this option represents one way of fitting together one's reassessment of the original examples and one's general views about ethical theorizing.

What I *do* propose to undertake here is a consideration of one important worry about the cogency of the method of argumentation employed above. If this worry is controlling, then it would provide (part of) a defense of the position that the Trolley Problem is a legitimate problem. The worry I have in mind is that the method of intermediation employed above may be illegitimate insofar as it is relevantly similar to the obviously specious forms of reasoning employed in *sorites* puzzles. Now it may be controversial and unclear exactly what mistake the proponent of sorites argumentation is making, but there is a general consensus that such a theorist *is* making *some* mistake. If so, and if the method of intermediation is relevantly similar to the sorites kind of argumentation, then one could defend the presuppositions of the Trolley Problem against the attack mounted above.

Briefly put, the sorites puzzle is a traditional problem in metaphysics. The name comes from the Greek word for "heap." The puzzle concerns when precisely a heap ceases to be a heap. That is to say, if you have a heap of sand and you take one grain away, it seems intuitive that you still have a heap of sand—how could taking one grain away make a heap into something that is not a heap any more? But presumably if you start with a heap of sand and continue taking grains away, at some point you will not have a heap any more. At what point does the heap cease to be a heap? How could taking away one grain from a heap transform it into a "non-heap"? These are some of the problems involved in the sorites puzzle. (Of course, the same puzzle can be applied to other sorts of entities.)

It must be admitted that there is a *prima facie* problem here. After all, I have (in section III) strung together various cases, claiming that, since case *A* is morally equivalent to case *B* and case *B* is morally equivalent to case *C*, then case *A* is morally equivalent to case *C*. (Sometimes the string includes more than three cases.) Is this sort of argumentation illegitimate? Is it relevantly similar to the specious kind of sorites argumentation?

I shall begin by giving some examples of arguments which appear to be paradigmatic examples of problematic sorites arguments. I shall give a metaphysical sorites argument and also two moral sorites arguments.

Having presented these arguments, I shall develop some considerations by reference to which it can be seen that the method of intermediation used above is importantly different from the sorites arguments.

Let us consider very briefly a metaphysical sorites. Think about any old swizzle stick. (Of course, nothing depends on the choice of this particular sort of object; it is simply an arbitrary.) If you take away one molecule, it is, presumably, still a swizzle stick. And, in general, it seems to be true that if an object with *N* molecules is a swizzle stick, then that object would still be a swizzle stick, if you took away one molecule. But now we can obviously generate the (mortifying) result that an object with, say, one molecule is a swizzle stick. (One is tempted to say that it would have to be a very small martini indeed!)

The result in question emanates from two claims: an existence claim and an assimilation claim. The existence claim is really a "possible existence claim." It claims that it is possible that there should be a certain sort of object—let us say a swizzle stick of *N* molecules. Now the assimilation claim here has the logical form of a universally quantified conditional: it says that, for any *N*, if you had a swizzle stick of *N* molecules and you took away one molecule, you would still have a swizzle stick. These ingredients generate the paradoxical result. More specifically, the paradox consists in the apparent plausibility of the premises (the existence claim and the assimilation claim), the soundness of the reasoning, and the manifest implausibility of the conclusion.

Now let us consider a few moral sorites arguments. First, it should be evident that the above sort of sorites reasoning can be applied to moral properties, such as "is a person." Think about any person of, say, *N* molecules. If you take away one molecule, the person is, presumably, still a person. And, in general, it seems to be true that if an object with *N* molecules is a person, then that object would still be a person if you took away one molecule. But now we can obviously generate the (alarming) result that an object with one molecule is a person. The reasoning is precisely parallel to the reasoning above, except that we are here dealing with a moral notion.

Here is another sorites argument in regard to a moral notion. It is, presumably, morally wrong to torture someone for ten hours. That is, given that you have no reason (other than that you want to see someone else suffer) to torture someone, it is morally wrong to do so. But, surely, if it is morally wrong to torture someone for ten hours, it is also morally wrong to torture someone for nine hours, fifty-nine minutes, and fifty-nine seconds. In general, if it is morally wrong to torture someone for *N* seconds, it is morally wrong to torture someone for *N-1* seconds. But it is an obvious implication of the above considerations that it is morally wrong to torture someone for zero seconds, (*i.e.*, that it is morally wrong not to torture someone at all) and this is clearly unacceptable. Again, as

above, the paradoxical result emanates from an existence claim—that there is a (possible) case of torturing someone for N seconds which is impermissible—and a universally quantified assimilation claim—that if it is morally wrong to torture someone for N seconds, it is morally wrong to torture someone for $N\text{-}1$ seconds.

It is also possible to construct a moral sorites that issues in the conclusion (arrived at by a different route than above) that it is permissible to save the five in "Transplant." We start with the claim that it would be morally permissible to extract the relevant organs (painlessly) *one second* before the individual would otherwise have died. Now the relevant assimilation thesis is: for any N, if it is morally permissible painlessly to extract the relevant organs N seconds before someone dies, then it is morally permissible to extract the organs $N+1$ seconds before the individual dies. But now it is evident how we can get to the result that it is permissible to save the five in "Transplant." Of course, it was suggested above that this claim is worth taking seriously; but it is in any case evident that the current *route* to the conclusion is entirely unacceptable. Having presented these examples of sorites arguments—metaphysical and moral—I am now able to state the worry about the method of intermediation rather crisply: how is the method of intermediation different from the form of argumentation involved in the clearly problematic sorites arguments?

B.

The answer to the worry can also be stated crisply, although it is a rather more delicate matter to give explicit content to the answer. I shall state in a somewhat abstract way what I believe the answer is, and then I shall attempt to give more precise content to it. I believe that, even if the specific way in which I undertake to explicate the answer is not entirely satisfactory, the broad outlines of the answer will still be correct. The specious sorites arguments all exploit the phenomenon of vagueness to derive their conclusions, whereas the method of intermediation does not. This is, I believe, the difference between the sorites arguments and the method of intermediation.

Recall the metaphysical sorites pertaining to the swizzle stick. The argument exploits the vagueness of the boundaries of a particular swizzle stick. Intuitively, swizzle sticks are such that it is not the case that one can have a swizzle stick and then take away one molecule and thereby create an object which is not a swizzle stick. In contrast, if the boundaries of swizzle sticks were not vague, there would be a determinate, precise point (in the sequence of subtraction of molecules) at which objects went from being swizzle sticks to not being swizzle sticks. It is this vagueness which evidently underwrites the assimilation thesis, which (as noted above) has the form of a universally quantified conditional. Simi-

larly, in the examples of moral sorites adduced above, the phenomenon of vagueness is exploited in certain ways.

Note that in all of the sorites arguments—metaphysical and moral—the assimilation theses are universally quantified conditionals. This is a *symptom* of the fact that they all exploit the phenomenon of vagueness. If a given notion involves vagueness, then there is some relevant continuum along which there is *no* point that is a definite boundary or cut-off point. Basically, this is what the universally quantified assimilation theses express. In contrast, the assimilation theses in the arguments presented above that employ the method of intermediation are *not* universally quantified conditionals; they are (or can be understood as) identity statements linking a few (two, three, four) cases. This is an indication that the method of intermediation is *not* exploiting the phenomenon of vagueness. The logical form of the assimilation theses in the sorites cases and the arguments employing the method of intermediation are crucially different, and this difference in logical form points to an important difference in the nature of the arguments: whereas the sorites arguments exploit aspects of the phenomenon of vagueness in an illegitimate fashion, the method of intermediation in no way depends upon vagueness. (This is *not* to say that the subject matter of the claims in the arguments employing the method of intermediation does not *admit of* vagueness. Rather, the point is that the method of intermediation does not illicitly fix upon and exploit this vagueness to generate its results.)

An alternative way of making the point that the two types of arguments are importantly different is as follows. The pattern of argumentation I have used above—the method of intermediation—does *not* have the surface form of the specious sort of sorites pattern. In the sorites pattern, one typically finds an inductive clause of roughly the following kind: if entity E has some property P, then some entity suitably related to E, $E+1$, has P. The paradoxical results are then generated by an extremely large number of applications of this inductive clause to some basic case. (The inductive clause is what I above called the assimilation claim, and the basic case was posited by what I called the existence claim.) But the form of the method of intermediation is fundamentally different; there is not a very large number of applications of some inductive clause specifying some very small change. Rather, three or four cases are alleged to be identical with regard to one feature of their moral status (although admittedly factually different in certain respects). In the sorites, there is a huge number of applications of some tiny change, and this (together with the fact of a certain sort of continuity) is exploited to generate the paradoxical results; in the arguments employing the method of intermediation, a few cases which are substantially different factually (*i.e.*, which do not differ only in some tiny way) are alleged to be morally similar in the relevant respects.

C.

I have just presented what I take to be the broad outlines of an answer to the question of what distinguishes the method of argumentation I used above from the bad old sorites form of argumentation. I wish now to make some tentative and preliminary efforts toward giving the answer more specific content. There are good reasons to believe that all of the sorites arguments are *unsound*: although the reasoning employed is valid, it is *not* the case that both premises (the existence claim and the assimilation claim) are true. Specifically, there are good reasons to think that the assimilation claims fail to be true. In contrast, there are no such reasons to deny the truth of the assimilation claims in the arguments which employ the method of intermediation. (In the final section of this paper, I consider the possibility that there are certain *other* reasons to deny the assimilation claims embodied in the arguments based on the method of intermediation.)

I shall now outline one way of solving the sorites puzzles.[11] I emphasize that this is the barest sketch of a solution, which would have to be filled in and defended against various objections if it were to be considered a serious candidate for a solution to the sorites puzzle. Here it suffices to lay out the main features of the approach, and to show how it would allegedly both solve the sorites puzzles and *not* similarly debunk the method of intermediation.

Let us start with the metaphysical sorites pertaining to the swizzle stick. Intuitively, the sorites illicitly exploits the vagueness of the boundaries of swizzle sticks—the fact that swizzle sticks do not have precise cut-off points. Given a particular swizzle stick, the vagueness of the boundaries of swizzle sticks implies that there are three classes of molecules: those molecules which are clearly inside the swizzle stick, those clearly outside the swizzle stick, and those of which it is neither true nor false that they are inside (or outside) the swizzle stick. The third class comprises a fuzzy penumbra of the swizzle stick: it is not clear whether any one of the molecules in the penumbra is inside or outside the swizzle stick.

In order to talk about such penumbral regions, it has been supposed (by some) that we need a three-valued logic or semantics.[12] On a three-valued logic, there would be the values, "True," "False," and "Indeterminate." Further, according to a plausible semantics for such a logic, a universally quantified statement would be true only if *all* of its instantiations would be true; thus, if some instantiations of a universal generaliza-

11. In presenting this approach, I am heavily indebted to ideas in Michael Tye, "Vague Objects," *Mind* XCIX (October 1990): 535-57.

12. For further discussion of this issue and related issues, see Tye, "Vague Objects." See also Kit Fine, "Vagueness, Truth, and Logic," *Synthese* (April/May 1975): 265-300.

tion are false or indeterminate, the universal generalization fails to be true.

On such an approach—which of course would need to be filled in and explained more thoroughly—it can be argued that the sorites problem of the swizzle stick is unsound. This is because its assimilation thesis fails to be true. Recall that this assimilation thesis is a universally quantified conditional of roughly this form: "For any number *N*, if you have a swizzle stick with *N* molecules and you take one molecule away, then you still have a swizzle stick."[13] But remember that there is a class of molecules of which it is neither true nor false that they are in the swizzle stick. Given this fact, there will be numbers such that if you have an object with those numbers of molecules, it will be *indeterminate* whether it is a swizzle stick. Further, there will be at least one such number, *N**, such that it is indeterminate whether an object with *N** molecules is a swizzle stick and also indeterminate whether an object with N*-1 molecules is a swizzle stick. Given these facts, the universally quantified conditional has at least one instantiation in which both its antecedent and consequent are indeterminate. Under such circumstances, it is plausible to ascribe the truth value, "Indeterminate," to the instantiation of the conditional. Thus, it is *not* the case that *all* of the instantiations of the generalization are true. It follows that the universal generalization (which is the assimilation thesis) fails to be true. And if so, the argument (which proceeds from the existence claim and the assimilation claim) is not sound.

Precisely the same considerations apply to the moral sorites pertaining to the notion of "person." Because of the vagueness of persons, there is a penumbra of molecules associated with persons: this penumbra is the set of molecules which are not definitely "inside" and not definitely "outside" of the person. Thus, there will be some numbers *N** such that it will be indeterminate whether an object with *N** (or *N*-1*) molecules is indeed a person. Thus, the relevant instantiations of the universal generalization will be indeterminate, and the universal generalization itself (which constitutes the assimilation claim) will not be true. Thus, the argument is not sound.

I have claimed that all of the sorites arguments exploit the phenomenon of vagueness in illicit ways. Further, I have shown how a certain strategy of response can be applied to both the metaphysical sorites argument pertaining to the swizzle stick and the moral sorites pertaining to the notion of personhood. I would claim that the same strategy works in the case of the sorites having to do with organ removal. I believe that a

13. I employ here a bit of technical terminology from quantification theory (or first-order predicate calculus). A "universally quantified" statement is of the form "For all x, F(x)." An instantiation of this universal generalization would be of the form "F(b) for some specific b." See, for example, Donald Kalish and Richard Montague, *Logic: Techniques of Formal Reasoning*, 2nd ed. (New York: Harcourt Brace Jovonovich, 1980), 88, 99.

slightly different strategy can be applied to the moral sorites in regard to the impermissibility of torture. Both moral sorites exploit the apparent vagueness of the relevant notions: the boundaries of personhood and the sequence of episodes of different intervals of torture. But whereas the boundaries of a particular person are genuinely vague, the vagueness in the sequence of episodes of torture is, I believe, merely apparent. The argument exploits the *apparent* vagueness of the sequence, but to deem the sequence genuinely vague is to commit a certain sort of error.

The response to the sorites has two important elements. First, one must avoid what Derek Parfit has called a "mistake in moral mathematics"—"ignoring small effects."[14] One second less torture is *less bad* than one second more, even if by an extremely small amount. Also, it seems to me that the nature of the sequence of episodes of torture is importantly different from the nature of the boundaries of persons: it is *precise*. That is to say, although it is (*prima facie*) wrong to torture someone for even a very short time, say, a second, it is *not* (*prima facie*) morally wrong to refrain from torturing someone at all. If there really are moral differences that issue from small changes in the amount of time of infliction of pain and the sequence of episodes of such infliction of pain is precise in the way I have suggested above, then one can say that the universal generalization which constitutes the assimilation claim is *false*.[15]

Now it may be that I am wrong about what I have said concerning the logical structure of this sorites. And it may be that similar sorites could be constructed (employing genuinely vague sequences of infliction of pain). If so, I would simply fall back on the first strategy of response (according to which the assimilation thesis is indeterminate). I do, however, believe that there are two distinct types of moral sorites. Although they both involve universally quantified assimilation theses and thus exploit the apparent vagueness of certain phenomena, one set of moral phenomena is genuinely vague, while the other is only apparently so.

Although the two strategies of response to the two sorts of sorites are slightly different, they are clearly related. They both claim that the sorites arguments are not sound because the assimilation premises are not true. In the first instance, where there is genuine vagueness, the assimilation premise is not true insofar as it is indeterminate. In the second instance, the assimilation premise is not true because there is no genuine vagueness. To the extent that a sound argument must have true premises, in both instances the sorites arguments are shown not to be sound.

---

14. Derek Parfit, *Reasons and Persons* (Oxford: Clarendon Press, 1984), 75-82.

15. Note that it is not enough (to block the sorites) to identify the possibility of mistakes in moral mathematics which are based upon failing to remember that small quantities can make a difference. This is because there is presumably an analogous phenomenon with respect to "tallness." It is also necessary to posit that the sequence has definite boundaries.

To summarize, the various sorites arguments—metaphysical and moral—have a different logical form from the arguments employing the method of intermediation. This difference is a symptom of the fact that the sorites arguments illicitly exploit aspects of the phenomenon of vagueness, whereas the arguments employing the method of intermediation do not. Further, I have sketched (in a preliminary and admittedly epigrammatic fashion) a strategy of response to the sorites arguments. This method of response targets the crucial assimilation theses; these are appropriate targets insofar as it is precisely a feature of these claims (the fact that the assimilation theses are universally quantified conditionals) that allows the arguments to exploit vagueness in an unseemly fashion. Finally, my claim is that a similar response does *not* apply to the arguments employing the method of intermediation. The fact that the assimilation claims in the arguments employing the method of intermediation are of a different form points to the fact that they are importantly different. They do not attempt to exploit the phenomenon of vagueness, and they cannot be defeated via the strategy of response applied to the sorites arguments. Thus, it can be seen precisely how the two different *routes* to the conclusion that it is permissible to save the five in "Transplant" differ: whereas the sorites route is unsound, the route employing the method of intermediation is not.[16]

Above I have merely sketched a certain approach to responding to some of the sorites arguments. This approach involves positing a third semantic value, "Indeterminate." I certainly have not developed the approach in detail. Nor have I defended it against various objections. Ultimately, it may turn out that this approach is unacceptable.[17] Note, however, that this would not vitiate my claim that the sorites arguments have an importantly different structure from that of the arguments employing the method of intermediation. I suppose that even if this strategy of response to the sorites is found to be untenable, the fact that an initially plausible strategy of analysis implies that the sorites arguments are unsound but does not imply that the arguments employing the method of intermediation are unsound provides at least *some* reason to think that the two kinds of arguments are importantly different.[18]

16. The argumentation employing the method of intermediation results in certain moral assimilations, such as that of "Bystander" and "Transplant." To get the result that it is indeed permissible to save the five in "Transplant," one needs the further claim that it is permissible to save the five in "Bystander."

17. For an interesting recent discussion of the sorites problem, see Mark Heller, *The Ontology of Physical Objects* (Cambridge: Cambridge University Press, 1990), esp. 68-109.

18. Of course, if it could be shown that the reason that the strategy of analysis fails to resolve the sorites paradox—if in fact it does fail—is precisely some feature of it which implies that the sorites arguments are different in structure from the arguments employing the method of intermediation, then the reason (alluded to in the text) would have been overridden (or preempted). I am not, however, sanguine about this possibility.

V

Someone might grant all of what I have said and still claim that the assimilation theses embedded in the arguments employing the method of intermediation are logically problematic. Although these assimilation theses are not vulnerable to the objections to which the parallel theses in the sorites arguments are vulnerable, they might nevertheless be logically flawed in *different* ways. I turn now to a consideration of various versions of this objection.

Recall that the method of intermediation posits a number of cases which are alleged to be relevantly similar. They are supposed to be factually similar in such a way as to entail no difference with regard to the relevant permissibility claims. We are dealing, then, with a sequence of cases which are alleged to have two kinds of properties: underlying causal-physical properties and supervenient moral properties (concerning permissibility).

Now one objection to the cogency of the method of intermediation (and the assimilation theses which putatively issue from it) is as follows. It may be the case that properties at one level can be subsumed under a single higher-level property and nevertheless be different properties. So, for example, different particular shades of gray are all subsumed under "gray"; nevertheless, they are not identical shades. Similarly, it would clearly be a mistake to assimilate cases which are different factually, simply because they both have the same moral property or fall under the same moral judgments.

One should not say that two cases are factually identical simply because they can be seen to have the same moral property. So, even if "Bystander" and "Transplant" are morally equivalent, it would not follow that they are factually the same, just as it does not follow from two shades' being shades of gray that they are the *same* shade. Sameness at the supervenient level does not entail sameness at the underlying level. But, of course, these sorts of objectionable claims are *not* made by the proponent of the method of intermediation. It is *never* alleged that the relevant cases are factually identical; the claim is merely that they are factually similar in such a way as to allow no moral difference with regard to the relevant permissibility claims. Thus, two shades, although clearly different shades, may both be shades of *gray*.

But let us look more carefully at the particular sequence involving "Bystander," "Ramp," and "Fat Man." The argument was that "Bystander" and "Ramp" are morally on a par, "Ramp" and "Fat Man" are morally on a par, and thus "Bystander" and "Fat Man" are morally on a par. More specifically, the argument is that the first two cases share some causal-physical property (or set of properties) by virtue of which they are morally similar in the relevant respect, and that the same is true of the second two cases. Thus, the first and third cases share some

causal-physical property by virtue of which they are morally similar in the relevant respect.

This sort of argumentation can be fallacious in certain contexts. For example, consider the following situation. Objects can have different colors (say on different parts). So, object 1 and object 2 can share the color red. In virtue of this common underlying property, object 1 and object 2 share the supervenient property, "having the same color (on some part or other)." And object 2 and object 3 can share the color blue. Because of this common underlying property, object 2 and object 3 share the supervenient property, "having the same color (on some part or other)." But it is clear that the above facts do *not* imply that objects 1 and 3 have the property of sharing the same color on some part or other—they may have no parts which share colors.

The reasoning described above is clearly fallacious. But the reasoning involved in the argumentation employing the method of intermediation is different. Note that there is a *shift* in the underlying property under consideration when moving from the first pair of cases (object 1 and object 2) to the second pair of cases (object 2 and object 3). The shift is from red to blue. This shift is what causes the problem, and the problem is created in part by the fact that the relevant properties—blue and red—are "compossible" (i.e., jointly exemplifiable). That is, the relevant properties are "being part red" and "being part blue"; these are compossible, and are jointly exemplified by object 2. Given this, the shift issues in the problematic result.

In contrast, the method of intermediation involves *no* shift in underlying properties. The underlying property in the above argumentation which employs the method of intermediation (for example, in the sequence, "Bystander," "Ramp," "Fat Man") is something like "being physically such as not to allow any relevant moral difference." This property and its complement are *not* compossible, and there is no shift in properties when proceeding from the first pair of cases to the second. Thus, nothing similar to the objectionable form of reasoning involving blue and red can be going on in the argumentation employing the method of intermediation.

Let us now consider a final version of the worry about the assimilation theses involved in the method of intermediation. Consider a sequence of colors. More specifically, consider a continuous sequence of physical propensities to reflect light which underlies a supervenient sequence of colors. Let us further suppose that although the sequence of propensities is continuous, propensity *P1* is considerably different from propensity *P3*—so different that they clearly are instances of different colors. It may nevertheless be the case that propensity *P1* and propensity *P2* are both instances of the same supervenient color, say, blue. And it may also be the case that propensity *P2* and propensity *P3* are instances of the same

color, say green. Note that this configuration of claims requires that *P2* be such as to admit of two "descriptions" or supervenient properties—blue and green. It is, we might say, "blue-green." Note that these are *not* descriptions of parts of objects, but (putatively) equally accurate descriptions of the entire propensity (or the *entire* objects which manifest the propensity). Thus, by transitivity it would seem to follow that *P1* and *P3* are instances of the same color. But this is, by hypothesis, false.

The reasoning here is clearly fallacious. But again it is *not* the sort of reasoning involved in the above argumentation employing the method of intermediation. What is problematic in the fallacious reasoning just presented is not a shift in an underlying compossible property, but a shift in a *supervenient* compossible property. When one assimilates propensities *P1* and *P2*, one is thinking of them as both *blue*; and when one assimilates propensities *P2* and *P3*, one is thinking of them as both *green*. The fallacious move is made possible by the fact that both blue and green (allegedly) can supervene on the same underlying physical propensity to reflect light.

But notice that nothing like this can be going on in the argumentation employing the method of intermediation. In this argumentation, the relevant underlying property is something like "physically such as not to allow any moral difference," and the relevant supervenient properties are moral properties such as "permissible to save the five" or "impermissible to save the five." When one argues that "Bystander" is to be assimilated to "Ramp" and "Ramp" is to be assimilated to "Fat Man," one is *not* proceeding in the following (admittedly problematic) way. One is *not* saying first that in "Bystander" and "Ramp" it is *permissible* to save the five (and thus the cases are morally equivalent in the relevant respect), and second that in "Ramp" and "Fat Man" it is *impermissible* to save the five (and thus the cases are morally equivalent in the relevant respect), and finally (by transitivity) that "Bystander" and "Fat Man" are thus morally equivalent in the relevant respect. This would indeed render the reasoning parallel to the fallacious reasoning presented above.

Rather, one is simply sticking to the underlying causal-physical property of being such as not to allow any moral difference. One is saying that "Bystander" and "Ramp" are physically such as not to allow any moral difference, "Ramp" and "Fat Man" are physically such as not to allow any moral difference, and thus "Bystander" and "Fat Man" are physically such as not to allow any moral difference. Note further that the supervenient moral properties, "permissible to save the five," and "impermissible to save the five" are *not* compossible. (These properties of course have been understood to represent "all-things-considered" judgments, rather than mere *prima facie* judgments.) Thus, the argumentation employing the method of intermediation cannot be problematic in the way in which the reasoning involving blue-green is.

In this section I have undertaken to defend the plausibility of the reasoning employing the method of intermediation against certain claims that it involves logical errors. I do not think that it is logically deficient. It may well be that the reasoning involves false claims or is otherwise infelicitous; I have not addressed these issues. Rather, I have simply attempted to defend the logical integrity of the form of reasoning.

Let me be a bit more explicit about this point. Whenever one presents a string of analogies (as I have above), one is essentially claiming that the pairs of cases are relevantly similar (or perhaps identical) in some respect (or respects). A clear logical error (or at least quasi-logical error) would be to *shift* respects. On the other hand, one might conceivably be making another sort of error: it might be that one's claims about the alleged similarities or identities are *false*. I have primarily been concerned with arguing that there are no obvious logical (or quasi-logical) errors in the argumentation that employs the method of intermediation as presented above.

Someone could, I suppose, challenge particular claims I have made about the cases. This would not be to challenge the *logic* of the argumentation, but it would challenge the soundness of the arguments (and thus the truth of the conclusions). Note, however, the very implausible claims that would need to be made by the proponent of such a challenge. Such a theorist would need to deny the moral assimilation (as regards the relevant claims) of cases such as "Bystander" and "Ramp." Perhaps this theorist would say that it is permissible to save the five in "Bystander" but impermissible to save the five in "Ramp." Or perhaps this person would say that it is permissible to save the five in "Bystander" but *indeterminate* whether or not it is permissible to save the five in "Ramp." But given the causal-factual structure of the cases, it is highly implausible to make such claims.

Note, finally, that it is extremely reasonable to think that at least *some* forms of reasoning employing transitivity must be valid, even if the sorites form of argumentation is rejected; it certainly does not follow from the existence of fallacious kinds of sorites arguments that *any* argument employing transitivity is unacceptable. Surely, *some* arguments in *some* contexts which employ transitivity are not relevantly similar to sorites arguments.

VI

Consequentialists and hybrid theorists[19] have this in common: they believe that it is always permissible to act in such a way that one would maximize the good consequences of what one does for the community as

19. The term is Scheffler's: see Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Clarendon Press, 1982).

a whole. Consequentialists also believe that it is always *obligatory* so to act, whereas hybrid theorists deny this. That is to say, whereas a consequentialist argues that one ought always to maximize the good consequences of one's behavior for the community as a whole, the hybrid theorist holds that it is not *obligatory* always to maximize the good consequences of one's behavior; rather, he argues that, whereas it is always *permissible* to perform the maximizing action, it is sometimes permissible to pursue one's own projects and commitments at the expense of maximizing good consequences for the community as a whole. Thus, the theory is a "hybrid" of consequentialism and non-consequentialism (deontology): consequentialist considerations apply to the relevant permissibility claims, whereas non-consequentialist considerations are allowed to play a role in the obligation claims.

Call the claim which consequentialists and hybrid theorists share, the "Permissibility Claim." I believe that the pattern of argumentation presented above (which generates reasons to assimilate the original cases) lends some support to the Permissibility Claim. Of course, it does not in itself *establish* such a claim, since the force of the argumentation is, strictly speaking, to *assimilate* various cases. It is this assimilation *in conjunction with* the further claim that it seems to be permissible to save the five in such cases as "Bystander" that yields support for the Permissibility Claim. Indeed, what is required is rather more than the assimilation claim and this claim about "Bystander"; one needs the claim that it is more plausible that the apparent permissibility of saving the five in "Bystander" should transfer to the other cases (given the assimilation claim) than that the apparent impermissibility of saving the five in "Transplant" should transfer to the other cases. (I feel some sympathy for this latter proposition, but I do not know how to argue for it.)

Philosophers such as Samuel Scheffler have argued that, whereas one can give an account of the picture of rationality which underlies consequentialism and the hybrid theory, it is difficult to construct a parallel account for non-consequentialism. Thus, insofar as one thinks that a normative ethical theory ought to be supported by some view about rationality, one has some reason to prefer consequentialism and the hybrid theory to non-consequentialism. Also, these theories appear to be preferable to certain versions of non-consequentialist or deontological approaches insofar as they are relatively systematic and algorithmic rather than intuitionistic. It might seem, then, that on general theoretical grounds, the Permissibility Claim is, at least, attractive.

But the existence of the kinds of cases involved in the Trolley Problem has caused anxiety about rushing to embrace the Permissibility Claim. Judith Jarvis Thomson begins her classic piece by saying, "Morally

speaking it may matter a great deal how a death comes about. . ."[20] And many have felt that the kinds of cases involved in the Trolley Problem provide strong reasons to reject the Permissibility Claim and to embrace some sort of non-consequentialism, despite the neatness and theoretical appeal of the Permissibility Claim. That is, non-consequentialist (deontological) approaches are often motivated by appeal to graphic hypothetical examples (of the sort discussed above), in which it is alleged that it is intuitively clear that the consequentialist and hybrid prescriptions are objectionable.

The significance, then, of the pattern of argumentation presented above is to provide support for the Permissibility Claim by casting some doubt on what is often taken to be a major obstacle to the acceptance of the Permissibility Claim.[21] If what I have suggested is true, then the sorts of cases involved in the Trolley Problem do not provide decisive reason to reject the Permissibility Claim and to embrace non-consequentialism. Of course, I do not claim that the cases I have adduced provide decisive reason to accept the Permissibility Claim—only that they generate a challenge to the view that certain cases provide insuperable obstacles to acceptance of the Permissibility Claim. Thus, I believe that I have provided some reason to doubt the main source of support for deontological ethical theories and the main source of anxiety about consequentialist and hybrid approaches. If the claim that consequentialist and hybrid approaches are in danger of countenancing obviously morally repugnant behavior—behavior which can readily be distinguished intuitively from permissible behavior—is called into question, then the appeal of deontology is substantially vitiated, and consequentialist and hybrid approaches are rendered significantly more attractive.[22]

---

20. Thomson, "Killing, Letting Die, and the Trolley Problem," 204.

21. Of course, in order to provide strong support for the Permissibility Claim, one would have to explore many difficult and complicated issues. There are, admittedly, delicate and hard issues which are relevant to the use of hypothetical examples of certain sorts to test ethical theories. These issues are clearly beyond the scope of this paper. (There will be some discussion of these issues in Fischer and Ravizza, eds., forthcoming 1991.) It should be pointed out that the argumentation presented in this paper provides an *internal* critique of the conclusions that some draw from the Trolley Problem. That is, *if* the original examples provide reasonable tests of a normative ethical theory, then the examples adduced here also seem to provide such tests. It would appear to me to be arbitrary to claim that, whereas discussion of the original cases is fair game, discussion of the other examples is not.

22. I have discussed the above issues with members of the UCLA Law and Philosophy Discussion Group, especially David Copp. Also, I have discussed these issues with members of the Moral and Political Philosophy Society of Orange County, especially Gary Watson and David Estland. I am grateful to these people, and also my colleague, Alex Rosenberg, for their helpful comments. I am deeply indebted to Mark Ravizza for his detailed and trenchant comments, and for many illuminating conversations.