Consequentialism, Group Acts, and Trolleys

Joseph Mendola

Abstract: Multiple-Act Consequentialism specifies that one should only defect from a group act with good consequences if one can achieve better consequences by the defecting act alone than the entire group act achieves. It also holds that when different valuable group agents of which one is part specify roles which conflict, one should follow the role in the group act with more valuable consequences. This paper develops Multiple-Act Consequentialism as a solution to the Trolley Problem.

Its relentless pursuit of the good provides act-consequentialism with one sort of intuitive ethical rationale. But more indirect forms of consequentialism promise more intuitive normative implications, for instance the evil of even beneficent murders. I favor a middle way which combines the intuitive rationale of act-consequentialism and the intuitive normative implications of the best indirect forms. Multiple-Act Consequentialism or 'MAC' requires direct consequentialist evaluation of the options of group agents. It holds that one should only defect from a group act with good consequences if one can achieve better consequences by the defecting act alone than the entire group act achieves, and that when different beneficent group acts of which one is part specify roles which conflict, one should follow the role in the group act with better consequences. This paper develops MAC as a solution to the Trolley Problem. Section 1 concerns the relative advantages of direct and indirect consequentialisms. Section 2 develops MAC by a focus on competing conceptions of group agency. Section 3 applies MAC to the Trolley Problem.

1

Act-consequentialism prescribes that each individual agent in each situation act in *direct* accord with the proper ordering of options from worst to best, indeed in its simplest form that each always choose the best.[1] A basic normative principle is applied directly to evaluate individual acts from among individual options. But consequentialism might alternatively assess things other than individual acts by reference to their valuable consequences, for instance moral rules for groups or stable individual virtues, and then assess individual acts *indirectly*, by reference to those other things. Hence *indirect* consequentialism, which has many forms.

Indirect forms of consequentialism seem to promise more intuitive normative implications than direct act-consequentialism. But indirect consequentialism faces obvious objection. One classic objection is that the various indirect forms, for example rule consequentialism, have the same implications as, are extensionally equivalent to, act-consequentialism. This would imply that those indirect forms are no more properly intuitive than the act-consequentialism on which they are supposed to be an intuitive improvement.[2] The argument for extensional equivalence seems straightforward. The best and most beneficent rules seem of necessity to allow for exceptions to any general restrictions on behavior which they proffer, exceptions which allow the local maximization of the good.

But in fact this objection is mistaken. None of the standard forms of indirect consequentialism is extensionally equivalent to standard act-consequentialism. There are perhaps many reasons for this, but one is important to us. One way in which extensional equivalence fails is instructive, since it suggests that the intuitive normative advantages of indirect forms result from the effects of cooperative behavior.

This key objection to extensional equivalence is Allan Gibbards'.[3] Here is his case:

> Smith and Jones . . . [are] placed in separate isolation booths, so that the actions of one can have no influence at all on the actions of the other. . . . [A] red push-button [is] installed in each . . . booth. The only action of moral significance open to either man will be to hold his push-button down at 10:00 a.m., or to refrain from doing so. . . . If at 10:00 a.m. both are holding down their push-buttons, they receive cake and ice cream. . . . If only one of them is holding his push-button down, however, they both receive electric shocks. . . . If neither is holding his button down, nothing happens.[4]

This is how the case works: Notice that if at 10 a.m. Smith is not holding his button down, then it is best for Jones to not hold his down. But then the best act-consequentialist act for Jones is not that which would be best for each to cooperatively engage in, nor that which would be best for rules to prescribe to both, nor best for both to do in similar circumstances.[5] The coordinating effects of cooperative action block the alleged extensional equivalence of act-consequentialism and indirect forms.

Of course, some indirect forms of consequentialism, for instance character-based consequentialism, focus not on the acts of more than one agent but on the acts of a single individual over many temporally distinct choice situations. But notice that the phenomenon Gibbard notes has a temporal analogue. It is simply Castaneda's paradox of act-consequentialism seen from a different angle.[6] If among your options are conjunctive acts, for instance acts which take some time and require a temporal conjunction of two shorter acts, it may be that the conjunctive act has the best consequences of all temporally extended acts available, but neither conjunct has good consequences on its own. The conjunctive act may be in effect a cooperative action of two temporal periods of your life.

So this objection to indirect forms of consequentialism fails. But it fails in an instructive way. Almost everyone grants that indirect forms have more intuitive normative implications as

long as they are not extensionally equivalent to act-consequentialism. But Gibbard's counterexample to extensional equivalence suggests that it is in particular the coordination of action in cooperative behavior which generates those intuitive advantages for indirect forms. This perhaps reflects the traditional Kantian and contractarian insight that cooperative activity and the respect and reciprocity which support it undergird norms which forbid lies, murder, and injury.

Still, there is an important complication. In Gibbard's counterexample, one of the two individuals *does not* perform their component of the joint act which would have best consequences. If they *had* performed their component, act-consequentialism and indirect forms would require the same act of the second individual. Act-consequentialism and indirect forms are apparently extensionally equivalent in their recommendations for that second individual in that situation. This is relevant because we will eventually see that the best form of consequentialism cedes normative salience to actual forms of cooperative behavior on the part of others. In other words, in the important cases of extensional equivalence or inequivalence for the view to be developed here, the relevant analog of Gibbard's first individual does his part in the cooperative scheme, while in Gibbard's case he does not.

But there are also other sorts of counterexamples, which reflect other aspects of cooperative activity, and which show a failure of extensional equivalence even in the cases most relevant to the view to be developed here. In the familiar Prisoner's Dilemma, there are two individuals who are so positioned that, if each acts directly to pursue their own self-interest, then each will do better in that regard whatever the other does, and yet both will lose relative to an outcome which was available by joint action. Act-selfishness and more indirect forms of selfishness are not extensionally equivalent even when the other prisoner in fact acts in a

cooperative manner. This has the structure of the case we need. And there are also moralized versions of the Prisoner's Dilemma, in which two consequentialists are positioned such that, if each acts directly to pursue best consequences, then each will do better whatever the other does, and yet both will lose relative to an outcome which was available by joint action.

It might seem that this could not be, since the two consequentialists, unlike the two selfish prisoners, share a goal. But it can be, for instance because the options available to someone depend on how they choose among their options. If someone chooses among options as an act-consequentialist, a Martian may torture all humans. If they choose as a deontologist or otherwise, it may promote the general welfare of all. Hence all the options of our consequentialist might be worse on consequentialist grounds if they choose as an act-consequentialist rather than not. Such a Martian can also assure the characteristic payoff matrix of a moralized Prisoner's Dilemma.[7] Of course, there are no Martians with nasty schemes like that one. But a more realistic counterexample is available whenever acceptance of certain traditional constraints against lying and murder makes better options available than would be available in a world of traditional act-consequentialists.[8] The consequent reduction in anxiety about the possibility of being murdered or injured for the common good would alone sometimes be sufficient. And truth-telling allows a kind of cooperative planning in the face of our ordinary ignorance of what others believe and will do under various conditions that is crucial to successful forms of cooperative behavior. And of course in the real world not everyone is an act-consequentialist. Others may refuse to admit someone who is an act-consequentialist into groups with beneficent goals, or may not be able to sufficiently trust an act-consequentialist to allow for the successful cooperative pursuit of some important consequentialist goal. Effective joint action also seems sometimes to require that some

of the cooperators abandon independent pursuit of the goal, even where trust is not in question.

And Gibbard also introduced another class of counterexamples to extensional equivalence of the sort we need, which indeed invoke other phenomena which are crucial for MAC. He calls these cases of 'surplus cooperation'.[9] Imagine that we all cooperate in a very beneficial practice of truth-telling. It would be very bad if many of us often lied, which would undercut the practice. But given the fact that almost no one ever lies, it may be beneficial for me to tell a few lies. I won't tell enough lies to undercut the practice with its significant general benefits, but I will grab a little extra utility on the side. There are *some* indirect forms of consequentialism which require that I not lie in this circumstance, while act-consequentialism suggests that I should. Indeed, what are probably the most intuitive indirect forms of consequentialism fall within the relevant space of this counterexample. In the instances most relevant to our eventual concerns, act-consequentialism and indirect forms are not extensionally equivalent for this second reason.

Indirect forms of consequentialism are not extensionally equivalent to standard forms of act-consequentialism. Therein lie their intuitive advantages. But there is also a second objection to indirect forms of consequentialism. They seem to many to have no coherent normative rationale. They seem to be ad hoc fixes which consequentialists adopt only because they deliver acceptably intuitive implications. If maximization of the good is the ultimate rationale of rule-consequentialism as of all forms of consequentialism, then it seems ad hoc or incoherent to suggest at the same time that individual acts which maximize the good are sometimes wrong because they violate rules general acceptance of which would have good consequences. It is a fact that consequences will sometimes be better served if an individual consequentialist defects from a co-operative consequentialist scheme and leaves its success for others to assure. And an

indirect form of consequentialism which assesses such a defecting act only via its fit with that best cooperative scheme seems to lack a coherent consequentialist rationale. While indirect forms of consequentialism are not extensionally equivalent to familiar direct forms, this characteristic intuitive advantage is apparently bought at the price of an incoherent rationale.

It may seem that there can be no middle ground. But there is. Multiple-Act Consequentialism enjoys at once the intuitive general advantage of indirect forms suggested by Gibbard's counterexample, which rest on the coordinating effects of cooperating activity, and also has a coherent and straightforward consequentialist rationale. The key to this possibility is the fact that the cooperative practices which the best indirect forms require are in fact group acts. Hence the ideal indirect form of consequentialism is extensionally equivalent to an unusual direct form.

Group action can be a conjunction of individual acts. And yet group action in accord with a consequentialist principle can have a direct consequentialist rationale. Some argue that indirect consequentialisms, which are supposed to have intuitive normative implications, in fact collapse into unintuitive act-consequentialism. This we have seen to be wrong. Rather, a new form of consequentialism, Multiple-Act Consequentialism, collapses happily into extensional equivalence with intuitive forms of indirect consequentialism. And it retains its coherent consequentialist rationale when it does so.

2

But what is Multiple-Act Consequentialism? This section develops Multiple-Act Consequentialism as an obvious implication of one natural conception of group action.

There are, I claim, group acts, performed by group agents. But there are group acts in a

specific sense.

The basic cells of group agents in my sense are what I will call atomic agents.[10] These are more or less momentary human agents, which persist for but a short time. They are brief time-slices of people, which have correspondingly short sets of options.[11]

Of course, momentary agents are not the most familiar agents. Familiar agents persist over whole lives. But this analytic convenience is not really revisionary. It makes familiar sense to talk of how someone's options change over time, or how they change their preferences or choices among stable options. Clearly quite brief periods of someone's life are sufficient to try for certain options over others, and even succeed at certain things, and hence can constitute agents with options.[12] That is all that the notion of an atomic agent requires.

But perhaps the notions of group agents and group acts seem more problematic. Such notions may seem mere metaphors. But notice that if we grant the existence of atomic agents, we should also grant that temporally persisting individual agents, the most familiar and obvious kind of agent, are a type of group agent. Remember Castaneda's paradox. Some intuitive individual actions require a temporally extended series of steps, which is one type of group action which requires extended temporal coordination among temporally distinct atomic agents. There is in such a case a single persisting group agent comprised of many cooperating atomic agents.[13] And common sense also recognizes the existence of intuitive group actions and agents of other kinds, involving several persisting humans. Certainly there are circumstances in which it makes vivid sense to talk of a group, say a gang or a university, trying one thing or another, or achieving one thing or another. This gives these groups alternative options in the crucial sense.

Even this very common speech may seem metaphoric, or to threaten some strange

monstrosity like absolute idealism. But literal group agency doesn't require anything troubling. To see this, and to more clearly refine the notion of group agents I will deploy here, it will be useful to consider the literature which develops more or less reductive analyses of group action and agents.

Let me be clear about my analytical claims here. While the notions of group action and group agent are familiar folk notions, I doubt that they are very determinate, specific, and clean-edged. And of course it wouldn't matter much if they were, as long as their edges were conventional or otherwise arbitrary. Still, it seems best to begin with whatever vague folk notions there are, and then work towards refining them to fit the needs of this project. To make our discussions more tractable, I will presume that group agency and group action go together. That does no obvious violence to the folk notions. In fact, I will begin a little further back than the relevant folk notions, with atomic agents. And then I will ask what collections of such agents constitute intuitive group agents capable of intuitive group action, allowing that some refinement of that intuitive notion might be needed or desirable.

Here is the core of my proposal, which as we will shortly see is a minor variant of Margaret Gilbert's conception: Group action and agency exists when there is common action by a number of atomic agents rooted in common true belief that there is a shared goal, and in acceptance by the members of the group that there is a reason to continue to coordinate activity until the goal is adequately accomplished, a reason whose acceptance we can expect to occasion criticism and the acceptance of criticism for failure to continue coordination until that point, or until the goal is mutually abandoned.

But let me add one small complication which will matter later on. It reflects Aristotle's

distinction between processes and actions, where 'processes' like building a house characteristically aim at a goal which extinguishes the activity in question, while pleasant 'actions' like surfing do not.[14] Let it rather be that the members of the relevant group accept that there is a reason to continue to coordinate activity until the goal is adequately accomplished, or indefinitely if it is not the sort of goal which is finitely and definitely accomplished, a reason whose acceptance we can expect to occasion criticism and the acceptance of criticism for failure to continue coordination until the point (if any) at which the goal is accomplished or mutually abandoned.

There are a variety of more or less reductive and yet richer accounts of group agency now in play among philosophers, and we can gain some illumination by a contrast of those alternatives with my somewhat stripped-down proposal.[15]

First, while my account is in fact closely modeled on Margaret Gilbert's, she deploys normative facts or real reasons where I deploy merely acceptance of reasons.[16] She suggests that people engaged in group activity *are entitled* to rebuke one another if they fail to keep their activity entwined. I suggest rather that they *accept that there are reasons* to rebuke one another under those circumstances.

The incorporation of normative facts seems an unfortunate feature of Gilbert's account, at least if we are to rest normative evaluations on MAC. But it is also apparently unnecessary to the central spirit of her proposal. What seems crucial is not the fact of an obligation, but rather merely the mutual acceptance that there is such an obligation. We can believe that there are witches when there are no witches, and we can accept that there are obligations when there are no obligations. As I would put it, to engage in group action its participants must accept that there is

a reason to continue to coordinate activity. This will serve to stabilize the activity, but without any genuine normative facts required. It will also explain the normative criticism and discussion we might expect in such a situation if one goes blithely out of coordination, and which Gilbert stresses as significant.

My modification of Gilbert's proposal will be helpful on other than merely metaphysical grounds. It will allow us to respect the obvious fact that there can be group action towards abhorrent ends, where people share a hideous project and hence have no true normative obligations of any sort to continue to coordinate on that project, whatever they may think and feel.

There is related but still more austere conception than my own which has some of these advantages over Gilbert's view, but is not really workable. For group action, I require acceptance of reasons, which has some motivational implications. But consider an alternative which requires that the members of a group merely believe that reasons exist without accepting them. The problem with this is that belief in such reasons is not enough. While genuine group action requires, I have claimed, not merely an attitude towards such reasons but also some pursuit by the group of its goal, still even if a department inadvertently stumbles towards some goal and everyone in the department recognizes the existence of some reason to continue cooperative activity towards that goal yet without accepting that reason, there is no intuitive group action. If everyone in the department cognitively admits the existence of reasons to criticize one another and to accept such criticism for failure to continue cooperative activity towards the goal, but no one is at all motivated to make that criticism or accept it, then in fact the department isn't trying.

It is also useful to contrast my proposal with richer analyses of group action besides

Gilbert's. Bratman makes this proposal:

> Shared cooperative activity (SCA) involves, of course, appropriate behaviors. . . . Given appropriate behaviors, what else is needed . . . ? Suppose that you and I sing a duet together, and that this is an SCA. I will be trying to be responsive to your intentions and actions, knowing that you will be trying to be responsive to my intentions and actions. This mutual responsiveness will be in the pursuit of the goal we each have, namely, our singing the duet. You may have this goal for different reasons than I do; but at the least we will each have this as a goal. Finally, I will not merely stand back and allow you to sing your part of the duet. If I believe that you need my help I will provide it if I can.[17]

There are a pair of relatively subtle elements of this view which are enrichments of the view I prefer, enrichments to which we will shortly return. But in any case Bratman's analysis of the remaining conditions undergoes some refinement until it becomes roughly this: Each of our paired singers sings their part, intends that they sing a duet, intends that they so act in accordance with and because of meshing subplans for action, and this is common knowledge between them. The stability of that intention in each individual, as suggested in Bratman's other work on the notion of intention as a plan, provides a certain stability of interaction and ensures certain sorts of mutual support.[18]

The relationship of this element of Bratman's account and my model turns on the relationship of stable intentions and accepted reasons. And it seems that in the central cases for humans these things do not come apart. There may be possible animals or even possible humans who have intentions but don't accept reasons. But, as Bratman himself suggests, you and I seem to treat our continuing stable intentions as reasons of at least some weight.[19] And if we were not so to treat them, then they would be insufficiently stable to constitute stable group activity. So while my variant of Gilbert's view deploys a slightly different basic notion than the weakened Bratman view immediately under consideration, there is at least a rough confluence between

those two accounts.

Nevertheless, there are the controversial enrichments which Bratman incorporates. The elements developed in the two preceding paragraphs are not enough, Bratman believes, for an SCA. Here's the first reason:

> You and I are singing the duet. I fully expect you to get your notes right, and so I intend to coordinate my notes with yours so that we sing the duet. But I have no disposition at all to help you should you stumble on your notes; for I would prefer your failure to our success. Were you unexpectedly to stumble I would gleefully allow you to be embarrassed in front of the audience. . . . And you have a similar attitude. . . .[O]ur singing may be *jointly intentional*; but it is not a SCA.[20]

But this enrichment doesn't seem intuitively necessary for genuine group action. Perhaps I am just not as nice a guy as Bratman. But I can easily imagine conscientiously engaging in a group action whose goal I favor, and yet hoping secretly that it will all fall apart because you will let down your end while I have kept mine up. Perhaps your embarrassment of that sort is a more significant goal to me than the goal of the group action, even though not something I will actively pursue, indeed not something that I *can* actively pursue. And even without such a secret goal, I can imagine being prepared to do my part in a group action, but not being prepared to go beyond my part to bail you out should you stumble in certain ways. I accept the reasons which underlie the group activity and grant them significant weight, but not as much weight, or perhaps not the kind of weight, which Bratman requires.

There is another enrichment present in Bratman's account. He requires that continuing action be mutually responsive.

> Suppose, for example, you and I lay plans for you to go to San Francisco while I go to New York. We might have a web of intentions concerning this joint activity, a web which satisfies . . . [the preceding conditions.] And our activity of prior planning may itself be a

> SCA. But if when we each go our separate ways there is no mutual responsiveness in action, our activity is prepackaged cooperation, not SCA.[21]

But it seems intuitive that the prior planning Bratman notes, which for instance might have specified the separate trips as part of an assassination plot, is sufficient to constitute group activity in this case, as long as the individuals cleave to the plan. On the other hand, it probably wouldn't be enough if they were still in reach of one another and failed to respond together in minimal and uncontroversial ways to changes in conditions. But this seems assured by the conditions required for group agency in my sense.

The notion of an SCA is Bratman's to develop as he wishes. But the two enrichments which Bratman suggests would not be conducive to the theoretical role in which we will deploy the notion of group action here.

Bratman's notion of an SCA is a little richer than the notion we need in two ways, as is Gilbert's in one. And both alternative accounts overlap more or less in the notion I will deploy. Now consider one more contrast.

Tuomela's account of group action is quite complicated. It is in some ways quite close to my proposal. But one of the complexities of Tuomela's view is that it is formulated in such a way as to suggest that there are more and less central cases of group agency, and the core cases are a bit richer than those I treat as crucial. Social group action, Tuomela believes, requires explicit or implicit agreement, and it requires that the group have an authority system, which determines the way in which individual wills determine group will.[22] In the paradigm cases of social group action, according to Tuomela, there is explicit agreement and also a relatively rich authority 'system', a characteristic way by which group members transfer their authority "over some issues

to the group, and the group . . . use[s] special operative members to form its will or . . . form[s] the group will through negotiation, bargaining, or voting with all its members acting as operative ones."[23] While an informal group carrying a table counts as a social group engaged in group action on Tuomela's view, it is only in a peripheral and somewhat degenerate way. It is only because that group can be conceived as having an implicit agreement and a degenerate authority system. But I believe that these complexities are unhelpful, at least for our purposes. Many of the group agents we will need to consider involve neither agreements nor authority systems of any reasonably concrete and robust sort.

But still, even the central group agents and actions in Tuomela's sense are simply enriched in a fourth direction from the core which my notion specifies. When an authority system is not accepted by its participants as generating reasons, when so-called implicit agreements are not recognized as reasons by the parties involved, the authority system is unstable and perhaps illegitimate, and there really is no implicit agreement.

We have seen that my notion of group action and agency captures one notion in which Tuomela's, Bratman's, and Gilbert's conceptions overlap, and which can be enriched in various different ways to get the more specific forms of group agency which they favor as an analytical focus. That seems an advantage of my account.

There aren't merely richer notions in the rough vicinity of my notion of group agency, but weaker notions as well. We have already considered an alternative which deploys belief in reasons rather than acceptance of reasons. But there are other possibilities. Consider packs of hunting animals who engage in cooperative and mutually responsive activity which isn't stabilized by anything closely resembling accepted reasons. Surely that is group action in some

intuitive sense.

But on the other hand it seems a form which is uncommon in humans when accepted reasons pull in another direction, and accepted reasons with us always seem to pull somewhere. While even humans can engage in some recognizable group activities--for instance certain sorts of rambunctious crowd response at concerts--which is not stabilized by accepted reasons and is indeed even such that all accepted reasons pull in another direction, still that seems no more to be intentional action by the group than an individual's trembling or fainting is intentional. The crowd in question is not trying to be raucous.

Group action in my sense is a fact. We have now located this phenomenon by reference to three richer and two weaker conceptions, and seen that it captures one reasonably folkish and apparently central notion of group action among at least humans. And we are now very close to Multiple-Act Consequentialism.

First, notice that acts in my sense are of intensely multiple sorts. There are atomic agents, but they are not the only kind. Some acts are parts of other acts, and not necessarily acts of the same agent. Some agents are parts of other agents. This means that there is a multiplicity not merely of numbers but of types of agents. Some agents, indeed the most intuitive, persist longer than atomic agents. There are various temporal scales of agents in the world--individuals in moments of their lives, in longer periods of their lives, or over whole lives. These are perhaps only somewhat unintuitively (though accurately) considered group agents. But some agents are even intuitive group agents, and of various types and scales. There are group acts of families, friendships, corporations, departments, and universities. All of these types of agents are real in my sense, and perform real actions from among real sets of optional alternatives.

Second, notice that this multiplicity is overlapping. Atomic agents in some sense constitute all agents. And particular atomic agents constitute at once parts of more than one group agent. You of this moment are at once part of a continuing person, and also various organizations, families, friendships, and the like. Indeed, even at a single moment you may be acting in some way which is relevant either by omission or commission to the group acts of all these groups. What you say now may at once support your career and an important family project but undercut your department and the crucial goal of your friendship with Y. In that way, it may at once violate your momentary role in certain group acts and play your momentary role in others.

All the multiple agents of which you are now part overlap in the atomic agent which you are at this moment. And atomic agents characteristically form parts of overlapping but real agents of different types. This overlapping multiplicity is reflected in the reasons accepted by atomic agents, which help constitute them as parts of more than one group agent all at once. You accept at this moment some reasons which help constitute the you of the moment as part of a persisting agent with persisting activities and plans, but also some reasons which help constitute you as part of a cooperating department, and others which help constitute you as part of a cooperating family.

We are now in position to understand the key normative tenets of MAC. According to Multiple-Act Consequentialism, atomic agents accept reasons which constitute, govern, and balance an overlapping multiplicity of genuine group agents. MAC specifies right acts for atomic agents, among other things. These acts are constituted as right by direct consequentialist evaluation of options. But this direct evaluation is performed all at once on the options of all the

multiple genuine agents who overlap in a given atomic agent. There may be, indeed there almost always will be, practical and hence normative conflicts among the various agents who overlap. MAC resolves these conflicts in two characteristic ways:

First, assume a situation in which an atomic agent can defect from a group agent with an appropriately beneficent project, and hence achieve additional good consequences on the side while not undercutting the project of the group agent. Now compare a first counterfactual situation in which the atomic agent achieves what it can by the defection but in which the various other atomic agents which constitute the group agent do not constitute such an agent, with a second counterfactual situation in which the group agent acts as it does and the atomic agent does not defect. If the first situation is better, then MAC says to defect. If the second situation is better, MAC says not to defect. This is the principle of Very Little Defection, or VLD.

What is the rationale behind this principle? It is the priority principle among actions suggested by a concern with maximization of the good. The acts are weighted by their consequences. You may object that the defecting agent can grab some extra benefit on the side, while the beneficent group act in fact persists. But while if we take the background of cooperative activity as fixed when figuring the options relevant to assessing defection, then defection gains a little and loses nothing, still if instead we take the action of the atomic agent as fixed, with the cooperative activity as the only variable item with normative weight, we get a predictably different answer. Since group acts are as real as individual acts, none has a natural priority. Rather we must let both relevant factors vary, and see which matters more on consequentialist grounds. And this principle seems the only workable way to do that.

And there is a second important principle with a similar rationale. Assume that we must

assess the relative importance of two forms of overlapping but conflicting group agency for an atomic agent which is part of both. Compare two counterfactual situations: In the first situation, the first group agent doesn't exist because the atomic agents in question fail to properly constitute such an agent, but the second group agent has its actual form. In the second situation, the second group agent doesn't exist, but the first has its actual form. If the first situation is better, then the second group agent is more normatively significant, and MAC says to defect from the first group agent in favor of the second. This is the principle of Defect to the Dominant, or DD.

There is more to be said about the rationale of these key normative tenets of MAC, and there are a variety of ways in which MAC can provide appropriate responses to standard objections to familiar forms of direct consequentialism. But here our focus will be on one specific intuitive application. We turn now to the interaction of MAC and the Trolley Problem.

3

One familiar objection to consequentialism is that it cannot deliver the commitments of commonsense morality to what Samuel Scheffler has called 'agent-centered restrictions'.[24] Agent-centered restrictions forbid particular agents from doing certain things, even if the cost of their failure to do that thing is that many other acts of exactly the same sort will be performed instead by others. For instance, an act which tortures the innocent seems intuitively forbidden. It may seem intuitively forbidden even if the certain cost of refraining from that act of torture will be that many other people commit many such tortures.

While MAC and agent-centered restrictions interact in a variety of ways, our focus here is more specific. Some object that the situation noted in the last paragraph is not coherently

described, because to fail to perform the first act of torture is in fact to commit the others. But common sense incorporates a series of normative distinctions between acting and refraining, doing and allowing, harming and letting be harmed, which may be deployed to defend this aspect of the familiar objection to consequentialism. The literature on the Trolley Problem treats in particular conflicts between intuitive agent-centered restrictions on killing and intuitive obligations to save lives, and reveals interesting and relevant difficulties in coherently resolving commonsense intuitions on such matters. As I said, that literature will be our focus here.

To get a sense of how MAC can help with the Trolley Problem, we will begin with a quick review of what it ideally should deliver. The literature on trolleys began with Philippa Foot's classic "The Problem of Abortion and the Doctrine of Double Effect".[25] But I will emphasize the also classic development of Foot's cases by Judith Jarvis Thomson in "Killing, Letting Die, and The Trolley Problem" and "The Trolley Problem".[26] There are a number of concrete examples to which we will need to attend to collect a set of relevant intuitions. The puzzles of the trolley literature arise from clanging together these diverse intuitions.

First, there is the case of Alfred and Bert. It is the standard assumption within the trolley literature that there isn't ALWAYS a morally relevant difference between otherwise identical cases of killing and letting die, and that this case shows that.

> [I]magine that
> (1) Alfred hates his wife and wants her dead. He puts cleaning fluid in her coffee, thereby killing her,
> and that
> (2) Bert hates his wife and wants her dead. She puts cleaning fluid in her coffee (being muddled, thinking it's cream). Bert happens to have the antidote to cleaning fluid, but he does not give it to her; he lets her die.[27]

Since these actions seem equally abhorrent, it seems that we should conclude with Thomson that

the difference between killing and letting die isn't always morally significant, which is of course not at all to conclude that it isn't sometimes quite significant.

Second, there is Foot's original Trolley case, and its slight but significant modification in Thomson's Bystander case. Here is Thomson's statement of the original Trolley:

> Suppose you are the driver of a trolley. The trolley rounds a bend, and there come into view ahead five track workmen, who have been repairing the track. . . . [Y]ou must stop the trolley if you are to avoid running the men down. . . . [T]he breaks don't work. . . . Now you suddenly see a spur of track leading off to the right. You can turn the trolley onto it, and thus save the five men. . . . Unfortunately, . . . there is one track workman on that spur of track. He can no more get off the track in time than the five can, so you will kill him if you turn the trolley onto him. Is it morally permissible for you to turn the trolley?[28]

With the majority, I intuit that this is morally permissible. Some think that this is a matter of choice between killing one and killing five, since as the driver of the trolley you are in charge of a hurtling object which will certainly kill someone. Some disagree. So, for clarity, let's introduce the slight modifications which give us the Bystander case:

> [Y]ou have been strolling by the trolley track, and you can see the situation at a glance: The driver saw the five on the track ahead, he stamped on the brakes, the brakes failed, so he fainted. What do you do? Well, here is the switch, which you can throw, thereby turning the trolley yourself. Of course you will kill one if you do.[29]

With Thomson and the majority, I intuit that it is morally permissible for you to turn the trolley. Hence sometimes it is morally permissible to kill one to save five. Indeed, though this is apparently not the consensus view, it seems to me to be morally mandatory in this case, though I would hesitate to blame someone who failed to turn the trolley, because of the moral difficulty of the situation.

But compare our third case, Transplant:

> [Y]ou transplant organs, and you are such a great surgeon that the organs you transplant

> always take. At the moment you have five patients who need organs. . . . If they do not get those organs today, they will all die; if you find organs for them today, they will all live. . . . [A] young man who has just come into your clinic for his yearly check-up has exactly the right blood-type. . . . Lo, you have a possible donor. All you need do is cut him up an distribute his parts among the five who need them. You ask, but he says . . . no. Would it be morally permissible for you to operate anyway?[30]

With the majority, I intuit that this would be morally impermissible. It is not always morally permissible to kill one even to save five.

To underline this fact, consider a fourth case, Fatman:

> [Y]ou are standing on a footbridge over the trolley track. You can see a trolley hurtling down the track, out of control. . . . [T]here are five workmen on the track. . . . [Y]ou know of one certain way to stop an out-of-control trolley: Drop a really heavy weight in its path. . . . It just so happens that standing next to you on the footbridge is a fat man, a really fat man. . . . [A]ll you need to do is give him a little shove.[31]

Here again I have the intuition, shared by almost all, that this is wrong.

So we have one case, Albert and Bert, which shows that sometimes there is no morally significant difference between killing and letting die, and a second case, Bystander, in which it seems at least permissible to kill one to save five. But we have two other cases, Transplant and Fatman, in which it seems impermissible to kill one even to save five.

Consider two more cases, which introduce other factors to which we will return. In reaction to the first new case, which I will call Bombing, many intuit that it is morally permissible to bomb an enemy factory while foreseeing but not intending that noncombatants will die, as a matter of collateral damage. But then consider Hospital Gas: "Suppose . . . that there are five patients in a hospital whose lives could be saved by the manufacture of a certain gas, but that this inevitably releases lethal fumes into the room of another patient whom for some reason we are unable to move."[32] In this case, with the majority, I intuit that this action would be

wrong, despite its abstract similarity to Bombing.

There is a large literature which attempts to develop general explanations which support our intuitive normative reactions to this range of cases. But none of the standard explanations generates intuitive results across the board. If negative rights are sometimes intuitively stronger than positive rights, still that difference doesn't seem relevant in the Bert and Albert case, and it apparently doesn't differentiate Bombing from Hospital Gas. If sometimes omissions which fail to save lives seem preferable to commissions which take lives, still one properly acts in the Bystander case to kill one, and that also doesn't explain the moral symmetry of Bert and Albert. If sometimes we have limited responsibility for foreseen but unintended consequences of our actions, that doesn't explain our reaction to Hospital Gas. If sometimes the temporal order of a harm and a causally connected good matters, that doesn't explain our differential reaction to Hospital Gas and Bombing.

Multiple-Act Consequentialism provides a new kind of general explanation of our various responses to these various cases. It is what might be called a framework explanation, which suggests the normative relevance of a variety of different factors in different cases. I believe it is in a sense a deeper level of explanation than those we have just canvassed, which explains why considerations such as those are important in some cases but not in others.

Still, it would be wrong to claim too much for MAC in this context. It is to a weak degree revisionary. It requires that not all our normative intuitions about cases be probative, though we will see that it does allow for the intuitive normative differences introduced through our survey of famous concrete cases. This slightly revisionary nature seems unobjectionable, since we *cannot* provide a coherent account which requires no revision at all in our intuitions about all

trolley cases. Why is this? For one thing, the order in which people are presented with concrete cases can affect the moral intuitions the cases generate. People seem to remember their initial response, and to try to preserve consistency in their responses.[33] Second, Unger has shown that cases which involve four options can yield an intuitive response which is plausibly inconsistent with the intuitive response to analogous two-option cases which involve merely the deletion of two options from the first.[34] Third, there are also other groups of cases which show together that something is sometimes wrong with our intuitions. Let me explain this third point.

Imagine a case like Fatman, but in which the trolley is merely slowed after running him over, and hence slowly comes to a stop before the five stuck on the tracks. Still, we retain our intuition that pushing him on the tracks to save the five is wrong. Now imagine instead that the fat man is tied on a side track, which runs back on to the main line before the point at which the five are stuck on the main track. Our choice as a bystander is between letting the trolley run on straight to kill the five, or turning it on to the side track, which will slow it down sufficiently when it runs over the fat man to save the five, even though the trolley will come back on to the main track and stop just before them. At least if that case is presented after the original modification of Fatman, most intuit that this is also wrong. Because we will run into this case again, I will give it a name, Sidetrack. Now imagine a third case, Loop. This is a lot like Sidetrack. The track which is on the side, with the fat man tied down, continues on beyond where he is tied to the tracks. But rather than coming back on to the main line immediately, as in Sidetrack, it instead loops around and comes back to the main line behind our five.[35] If we do nothing, the five will be run over directly. If we turn the trolley, it will be slowed by running over the fat man before it gets back to our five, though it will come at them from behind, and would

have run them over from that direction if it hadn't run him over. At least if presented this case after the preceding cases, most people intuit that it is wrong to turn the trolley on to the side track, because the difference in the way in which the track is arranged in Sidetrack and Loop seems trivial and irrelevant. But notice that Loop is quite like Bystander, at least if there is fat man on the side track in that case, but with the slight addition in Loop of some unused track connecting the track running beyond the fat man and the main line. It is hard to believe that unused track could be of grave moral significance, but in Bystander most intuit that it is permissible to turn the trolley. These cases together suggest that we cannot coherently treat all our intuitions about trolley cases as probative and stable.

Despite the fact that some revision of our intuitions about cases is inevitable and desirable, still Multiple-Act Consequentialism provides a mechanism to explain the intuitive differences among the various standard cases with which we began this section. In essence, MAC suggests a range of small factors which are sometimes relevant, and may sum together between various cases in various ways to deliver intuitive results. To develop this conception, first let me note some kinds of factors which MAC suggests may be relevant. Then I will go on and apply this analytical framework to deliver some of the more difficult intuitive differences among the standard cases with which we began.

Possible factors: First, when there is a particular group agent present in a case, that may cause a difference in the proper normative assessment of such a case according to MAC. For instance, it might allow us to deflect a risk within a cooperating group agent but not onto an outsider. Or, alternatively, it might be specially important to deflect the risk outside the group to defend group cooperation. Notice that people tied together on tracks might come for that very

reason to share accepted reasons that constitute them as a group agent of some sort. Or alternatively it might be that because they are tied down the individuals are not cooperating agents at all and hence not subject to the customary level of normative protection which that provides.

Second, it might be that the exact form of proper treatment of the deflection of risks and the like depends on a general cooperative practice about such things which is itself a group agent, a group agent which constitutes a general manner of handling the deflection of risks. If we all agree that people shouldn't be thrown on the tracks when in the position of the fat man in the original Fatman case, if we have a cooperative practice which forbids that, that will be normatively relevant according to MAC when the practice meets the conditions, noted in the last section, required for group agency.[36]

Third, temporal order is tied up with the existence of cooperative group agents in interesting ways. Perhaps only those who precede an event can cooperate to prevent it.

Fourth, there are obvious differences in the forms of group agency which link people who share a hospital as doctors and patients and those which link or fail to link different populations at war.

While as I've said, MAC provides a mere framework explanation of trolley cases, in which specific group acts suggest the importance of different sorts of factors in different cases, it certainly suggests the general overall importance of groups. That in itself may seem like another general explanation of our intuitions, on a par with the distinction between commission and omission or the Doctrine of Double Effect. And indeed both Robert Hanna and Peter Unger have suggested that at least many of our different intuitions about these concrete cases can be

generally explained by reference to our treating various groupings of people as normatively salient. We intuitively distinguish between innocent bystanders, like the fat man on the footbridge, and those already at risk in a situation, say those tied on the tracks in the same area. Unger thinks that this is a cognitive distortion.[37] But Hanna thinks that it is appropriate.[38] MAC suggests a kind of middle course. Certain groups, namely those constituting group agents or recognized as salient by specific group practices constituting group acts, are quite normatively salient, and may in different cases contribute different sorts of normatively relevant factors. But other intuitive groups may introduce cognitive distortions.

By way of contrast, consider Hanna's general account: There are moral situations to which some people are bystanders and in which some people are participants. It is impermissible for anyone to force a bystander to become a participant victim in a trolley case, but it is permissible for a participant to harm other participants to establish a better outcome.[39] The fat man on the footbridge is a bystander, but anyone already tied on relevant tracks is a participant. So it is okay to turn the trolley and save five at the cost of one, but not to push the fat man on the tracks.

I think that this is more or less right. Hanna's account of moral situations is expressed in a way which suggests rich and I think implausible metaphysical presumptions, for instance the denial of physicalism.[40] But I believe this to be an avoidable aspect of his presentation of the view. Still, while he provides hints about the distinction between bystanders and participants,[41] he does not provide an account of that distinction which would allow some independent rationale for our intuitive normative judgments about the cases. He rather suggests that we should work backwards from our intuitive normative judgments about cases to the distinction.[42] This is

problematic, because sometimes our intuitive judgments are affected by cognitive distortions. Remember the importance of the order of presentation of cases, and the other forms of incoherence in our judgments about cases noted above, which seem sometimes due to distortions introduced by grouping. What's more, it would be better to provide an account which rationalizes our probative intuitive judgments about such cases, not an account which merely summarizes our intuitive judgments.

MAC provides a criterion for the distinction between properly relevant participants and bystanders which provides a rationale for our probative intuitive reactions to cases, and which suggests that while certain intuitive groupings introduce cognitive distortions in the way that Unger suggests, many do not. It provides a kind of refinement of Hanna's general proposal, but which also invokes specific distinct factors of the sorts noted above, and which provides a reductive criterion for the bystander-participant distinction yielding a suitable rationale for a coherent and properly intuitive response to all trolley cases.

So far I have only made rough suggestions about a kind of mechanism which MAC can apply to trolley cases. What may be more telling is an application of MAC to deliver some particular commonsense intuitions about differences between particular cases. As I've said, unlike standard attempts at unified explanations of all our intuitions in trolley cases, MAC suggests that a variety of small factors which involve cooperative action in different specific ways can sum together to generate a significant difference in the moral status of different actions. But we need to see how this works in more detail.

Let's begin with cases which are intuitively far apart. Foot's original Trolley case and Transplant certainly occasion different intuitions. But notice that a trolley driver is a professional

who is plausibly specially responsible for the trajectory of their trolley by virtue of a specific cooperative practice. This may constitute a group action, underwriting a special professional obligation to turn the trolley in the direction where it will do least harm. Notice also, coming from the other direction, that the doctor in Transplant is also a professional member of a cooperative practice providing health care, indeed a number of overlapping or concentric cooperative practices of that sort constituting various group agents, who has special obligations to patients in the hospital, and in general not to cause death.

But those are not the only relevant moral differences between these two cases. To the end of identifying some of the other factors, consider intuitively intervening cases, for instance Bystander and Fatman. It seems morally permissible or even perhaps required to turn the trolley in Bystander, but morally impermissible to push the fat man onto the tracks. How can MAC deliver this difference?

We have, I think, a real intuition that supports the Pauline Principle, the principle that you should not do evil (like pushing the fat man in front of the trolley to his certain death) even if greater good would come of it (the life of the five), while on the other hand it is sometimes appropriate to do some great good while foreseeing that smaller evil will come of it (like turning the trolley in Bystander). We also have, I believe, a real intuition that you shouldn't drag an innocent like the fat man into a situation like the one in question against their will. We need to be a little careful, however, in how we handle these intuitions. For one thing, in real life situations, when you do evil as a means to the good, the temporal priority of the evil makes it more certain than it would be if the temporal order of the good and evil were reversed. It is hard to control for the effects of this fact when considering hypothetical cases. Second, I have already noted that not

all our intuitions about cases like these can be treated as probative. Still, there is something to the basic intuitions suggesting some normative weight for the Pauline Principle and for the distinction between doing evil and refraining from doing good. But MAC can deliver that weight.

Recall, in this order, Fatman, Sidetrack, Loop, and Bystander. We start with a case in which action to save the five seems intuitively quite wrong, and then end up with a case where it seems permissible and even maybe required. Here is an analysis of this phenomenon according to MAC:

Between Fatman and Sidetrack, these factors intervene: A) There is a basic scheme of group action of which almost all of us are participants, which is a kind of cooperative agreement not to draw uninvolved outsiders into such a situation against their will. If our generally accepted background practice were different, if for instance we had all agreed to share risks in some more rational but still intuitively very scary way, that would make this factor irrelevant. But under actual conditions it is relevant. It is very weighty by the basic normative tenets of MAC. B) There may also be a second set of factors which MAC can deliver, though I am less confident of this. The fat man in Fatman is outside the originally affected group. At least if the people on the tracks are conscious, they will share acceptance of reasons supporting group action to get at least some of them to safety. They are a group agent of one kind. Of course, on the other hand, the practical salience of this project may be minimized if they are tied down, and a person tied on a side track will only share a relatively abstract group project and accepted reason with those on the main track. But there is also another intuitively salient group which includes those who are stuck on the track and the choosing agent in the case. And the fat man will not share acceptance of the relevant reasons with that group. He is in a recognizable sense outside of a relevant group which

includes all of those on the tracks, and the other relevant group which includes those on the tracks and the choosing agent. And these groups correspond to certain group agents in my sense.

These two factors, or at very least the first, introduce some weak normative differences according to MAC. But weak differences are the most which commonsense intuition finds between these cases anyway. As we move from Sidetrack to Loop, all those on the track seem even more closely linked by these weak factors.

Between Loop and Bystander, there are various factors which interfere with the probity of our intuition, for instance the unreality of the very clear risks in trolley cases and also temporal prejudice. But at least the temporal factors might introduce a genuine difference accommodated by MAC, since almost all of us plausibly share a general scheme of group cooperation in accord with something like the Pauline Principle which plausibly constitutes a form of group action. We have a practice in which it is one thing to deflect a harm so it kills someone else and quite a different thing to deflect the harm by killing that someone. And that practice is a normatively weighty group act we share.

There are certain advantages according to MAC in the standing, general cooperative practices for distributing risks and harms on which my analysis has crucially relied. Alternatives are scary and unfamiliar. Still, these actual practices have costs also, and possible alternative practices might conceivably have better consequences. It is relevant according to MAC what cooperative practices of this sort we actually have, but they must also have basically beneficent results to carry normative weight according to MAC. And it is also relevant that there may be prospective and alternative forms of cooperation which would have better results. I believe that the background practices which give priority to leaving the fat man undisturbed are at least of

questionable morality according to MAC. But if they are proper, which after all common sense does presume, then MAC can deliver the intuitive differences in our reactions to Fatman and Bystander. Still, there is no deep difference here, but a series of small differences which are indeed dependent in certain ways on our contingent conventions.

We have seen a variety of small phenomena which intervene between Transplant and Bombing and between the other four cases which are in some intuitive sense between them. They create normative differences according to MAC, which track intuitive differences between the cases. There are other similar factors which MAC suggests may be relevant in other cases. For instance, we intuit that the presence of other human agents in a causal chain breaks our responsibility to a degree for the outcome of that chain, at least if they are not turned into mere instruments or means by threats or the like. But MAC might cede those other human agents a special normative role of that sort if there is a form of general group agency which makes it specially those others' business to interrupt such a chain. This might be assured by a general cooperative scheme like those for distributing risks and harms, or by the structure of a more particular and local group agent. To the degree that someone in the intervening chain fails to be a cooperating agent, this would undercut, according to MAC, the significance of their presence in the chain for the commutation of our responsibility for the relevant outcome.

MAC can deliver intuitive normative implications while preserving the natural normative rationale of act-consequentialism. In particular, it can provide a suitable treatment of trolley cases.[43]

Department of Philosophy

University of Nebraska--Lincoln

Notes

1. There is a distinction between subjective and objective consequentialist theories, which in their simplest direct act forms suggest respectively that the right act is best given our *conception* of our options or the *fact* of our options. I presume the latter.

2. For classic forms of this argument see David Lyons, *Forms and Limits of Utilitarianism* (Oxford: Clarendon Press, 1965), and R. B. Brandt, "Toward a Credible Form of Utilitarianism", in H. N. Castaneda and G. Nakhnikian (eds.), *Morality and the Language of Conduct* (Detroit: Wayne State University Press, 1963) pp. 107-144.

3. Allan F. Gibbard, "Rule-Utilitarianism: Merely an Illusory Alternative?", *Australasian Journal of Philosophy* 43 (1965) pp. 211-220.

4. Gibbard, "Rule-Utilitarianism" p. 214.

5. There is a similar example developed in Allan Gibbard, *Utilitarianism and Coordination* (New York: Garland Press, 1990) pp. 6-7. This is a reprint of his 1971 dissertation.

6. Hector-Neri Castaneda, "A Problem for Utilitarianism", *Analysis* 33 (1973) pp. 141-142. For discussion see Fred Feldman, "World Utilitarianism", in K. Lehrer (ed.), *Analysis and Metaphysics* (Dordrecht: Reidel, 1975) pp. 255-271, and *Doing the Best We Can* (Dordrecht: Reidel, 1986).

7. This case is a slight modification of a case in Joseph Mendola, "Parfit on Directly Collectively Self-Defeating Moral Theories", *Philosophical Studies* 50 (1986) pp. 153-166.

8. Gibbard, *Utilitarianism and Coordination* pp. 157-237 argues that in a society of act-utilitarians an established convention of agreement-keeping might under certain conditions create a coordination point sufficient to allow a practice of agreement-keeping among act-utilitarians, and that such a convention could be established among such utilitarians by a teaching practice of a certain sort described there. He also argues more generally that in an act utilitarian society "the agreements that would be kept include almost all the agreements to which act-utilitarians want to bind themselves." See p.159, though note also the qualifications pp. 201-205. On pp. 158-159, he further suggests that the argument extends to truth-telling, but does not explain the extension. Nevertheless, this does not undercut the extensional inequivalences noted in the text. Indeed, Gibbard himself suggests the second class of counterexamples we will discuss.

9. Gibbard, *Utilitarianism and Coordination* p. 20.

10. Thanks to Guy Rohrbaugh and Mark van Roojen for the phrase.

11. I will presume that options of an agent must be taken in the period that the agent exists, that things an agent does must be done during the period it exists. This is the commonsense view except perhaps for unusual cases. For instance, I may insult you after my death by a carefully

judged sentence in my will. But the presumption is put under strain by a focus on more or less momentary agents. The problem isn't that the options which I consider in the moment include longer acts than that; indeed that is part of my point in introducing the notion of atomic agents, to reveal the way in which atomic agents consider options of what are in fact group agents which constitute one persisting life. Rather the problem is that we don't often consciously intend or attempt actions which take but a moment to perform. But the options relevant to agents are things that they can do if they try *something*. You can hole a putt, but only if you focus your attention elsewhere, and yet, that is one of your relevant options. So the momentary options needn't be defined by tryings with contents focused in the moment. In fact, they characteristically aren't.

12. But see the previous note.

13. This is a theme of Christine M. Korsgaard, "Personal identity and the unity of agency: A Kantian response to Parfit", in *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996) pp. 362-397, and Carol Rovane, *The Bounds of Agency* (Princeton: Princeton University Press, 1998). If options and actions which someone can perform include conditions which obtain after their death, or after their agency ceases, then this conception of a continuing agent would need to be augmented in certain ways. But this should not be important to what follows.

14. Aristotle, *Nicomachean Ethics*, Book I, 1.

15. Margaret Gilbert, *On Social Facts* (London: Routledge, 1989); *Living Together* (Lanham, MD: Rowman & Littlefield, 1996); *Sociality and Responsibility* (Lanham, MD: Rowman & Littlefield, 2000). The essay "Walking Together", in *Living Together* pp. 177-194, is a good introduction to this complex project. Michael E. Bratman, "Shared Cooperative Activity", *The Philosophical Review* 101 (1992) pp. 327-341; "Shared Intention", *Ethics* 104 (1993) pp. 97-113. Raimo Tuomela, *The Importance of Us* (Stanford, CA: Stanford University Press, 1995).

16. Gilbert, "Walking Together" pp. 178-180.

17. Bratman, "Shared Cooperative Activity" pp. 327-328.

18. Michael Bratman, *Intentions, Plans, and Practical Reason* (Cambridge, MA: Harvard University Press, 1987).

19. Bratman, *Intentions, Plans, and Practical Reason* p. 51.

20. Bratman, "Shared Cooperative Activity" pp. 336-337.

21. Bratman, "Shared Cooperative Activity" pp. 339.

22. Tuomela, *The Importance of Us* pp. 171-227.

23. Tuomela, *The Importance of Us* p. 177.

24. Samuel Scheffler, *The Rejection of Consequentialism* (Oxford: Oxford University Press, 1982).

25. Philippa Foot, "The Problem of Abortion and the Doctrine of Double Effect", *Oxford Review* 5 (1967). This is reprinted in her *Virtues and Vices and Other Essays* (Berkeley, CA: University of California Press, 1978), and also in John Martin Fischer and Mark Ravizza (eds.), *Ethics: Problems and Principles* (Forth Worth, TX: Harcourt Brace Jovanovich, 1992) pp. 60-67.

26. Judith Jarvis Thomson, "Killing, Letting Die, and the Trolley Problem", *The Monist* 59 (1976) pp. 204-217; "The Trolley Problem", *The Yale Law Journal* 94 (1985) pp. 1395-1415. Both articles are reprinted in Fischer and Ravizza.

27. Thomson, "Killing, Letting Die, and the Trolley Problem", in Fischer and Ravizza, p. 69.

28. Thomson, "The Trolley Problem", in Fischer and Ravizza, p. 280.

29. Thomson, "The Trolley Problem", in Fischer and Ravizza, p. 281.

30. Thomson, "The Trolley Problem", in Fischer and Ravizza, pp. 280-281.

31. Thomson, "The Trolley Problem", in Fischer and Ravizza, p. 288.

32. Foot, "The Problem of Abortion and the Doctrine of Double Effect", in Fischer and Ravizza, p. 65.

33. Peter Unger, *Living High and Letting Die (*Oxford: Oxford University Press, 1996) p. 92.

34. Unger, *Living High and Letting Die* pp. 88-94. My guess is that this is due to the fact that cognitive processing of complex cases involves a certain sort of abstraction and simplification, at least when that effect isn't swamped by the order of presentation of cases.

35. Loops like this were introduced by Thomson. See "The Trolley Problem", in Fischer and Ravizza, p. 285.

36. John Harris, "The Survival Lottery", *Philosophy* 50 (1975) pp. 81-87, introduced the possibility of alternative risk sharing practices.

37. Unger, *Living High and Letting Die* pp. 96-101.

38. Robert Hanna, "Morality *De Re:* Reflections on the Trolley Problem", in Fischer and Ravizza, pp. 320-336.

39. Hanna, "Morality *De Re:* Reflections on the Trolley Problem", p. 330.

40. Hanna, "Morality *De Re:* Reflections on the Trolley Problem", p. 329.

41. Hanna, "Morality *De Re:* Reflections on the Trolley Problem", pp. 329-330.

42. Hanna, "Morality *De Re:* Reflections on the Trolley Problem", p. 333.

43. Thanks to Tom Carson, Robert Hanna, Guy Rohrbaugh, and Mark van Roojen for helpful comments on this paper.