

Does Participation Matter? An Inconsistency in Parfit's Moral Mathematics

BEN EGGLESTON

University of Kansas

Consequentialists typically think that the moral quality of one's conduct depends on the difference one makes. But consequentialists may also think that even if one is not making a difference, the moral quality of one's conduct can still be affected by whether one is *participating* (even if only ineffectually, or redundantly) in an endeavour that does make a difference. Derek Parfit discusses this issue – the moral significance of what I call 'participation' – in the chapter of *Reasons and Persons* that he devotes to what he calls 'moral mathematics'. In my paper, I expose an inconsistency in Parfit's discussion of moral mathematics by showing how it gives conflicting answers to the question of whether participation matters. I conclude by showing how an appreciation of Parfit's error sheds some light on consequentialist thought generally, and on the debate between act- and rule-consequentialists specifically.

I. INTRODUCTION

Woody Allen once said that ninety per cent of success is just showing up. But success is one thing; morality is another. Consequentialists, especially, may think that the moral quality of one's conduct depends on the *difference* one makes. Still, consequentialists may also think that even if one is not making a difference, the moral quality of one's conduct can be affected by whether one is *participating* (even if only ineffectually, or redundantly) in an endeavour that does make a difference. So consequentialists may think that what matters, morally, is not only making a difference of some kind, but also merely participating in making a difference – the moral equivalent of just showing up.

Derek Parfit discusses this issue – the moral significance of what I call 'participation' – in chapter 3 of *Reasons and Persons*.¹ This chapter is a discussion of what Parfit regards as five 'mistakes in moral mathematics'. The five 'mistakes' are, according to Parfit, misunderstandings that a person can have regarding what should be considered the consequences of an act, or regarding how the moral assessment of an act depends on its consequences. In identifying and correcting these five 'mistakes', then, Parfit aims to improve our moral assessments of acts by improving our accounting of the consequences of acts.

We will see that one of these 'mistakes', the second 'mistake', is essentially the principle that participation does not matter: the principle that merely participating in an endeavour is not the sort of act

¹ Derek Parfit, *Reasons and Persons*, Oxford, 1984.

that an agent can have moral reasons for performing or refraining from. Some people think that, far from being a mistake, this principle is one that consequentialists must embrace. Indeed, Frank Jackson has claimed this in a recent paper.² But as I have argued elsewhere,³ consequentialists may consistently say (as Parfit says in his indictment of the second ‘mistake’) that participation matters. What I argue in this paper is that *Parfit* cannot consistently say this, because of other things he says about moral mathematics. I argue that in his indictment of the *first* ‘mistake in moral mathematics’, Parfit implies that participation does *not* matter. The upshot is that in his indictments of the first and second ‘mistakes’, Parfit gives conflicting answers to the question of whether participation matters.

In order to reach this result, I devote sections II and III to a summary of the two ‘mistakes’ in question, along with Parfit’s arguments against them. In section IV, I make explicit the sense in which Parfit’s indictment of the second ‘mistake’ implies that participation matters, and in section V, I show how Parfit’s indictment of the first ‘mistake’ implies the opposite. After pausing in section VI to reflect briefly on Parfit’s contradictory passages, I devote section VII to an examination of another of Parfit’s passages that may seem to offer a resolution of this inconsistency. I conclude, in section VIII, by showing how an appreciation of Parfit’s error sheds some light on consequentialist thought generally, and on the debate between act- and rule-consequentialists specifically.

II. THE FIRST ‘MISTAKE’

The first ‘mistake’ is to take what Parfit calls the ‘Share-Of-The-Total View’, which he introduces through an example:

The First Rescue Mission: I know all of the following. A hundred miners are trapped in a shaft with flood-waters rising. These men can be brought to the surface in a lift raised by weights on long levers. If I and three other people go to stand on some platform, this will provide just enough weight to raise the lift, and will save the lives of these hundred men. If I do not join this rescue mission, I can go elsewhere and save, single-handedly, the lives of ten other people. There is a fifth potential rescuer. If I go elsewhere, this person will join the other three, and these four will save the hundred miners. (pp. 67–8)⁴

On the Share-of-the-Total View, if I stand on the platform with the first three of the other rescuers, then the consequence of my act is to save

² Frank Jackson, ‘Which Effects?’ in Jonathan Dancy (ed.), *Reading Parfit*, Oxford, 1997, pp. 42–53.

³ Ben Eggleston, ‘Should Consequentialists Make Parfit’s Second Mistake? A Refutation of Jackson’, *Australasian Journal of Philosophy*, lxxviii (2000).

⁴ Simple parenthetical references are to Parfit’s *Reasons and Persons*.

one fourth of the hundred lives, or twenty-five lives. And if I single-handedly save the other ten lives, then the consequence of my act is to save ten lives. So on this view, the consequence of my joining the other rescuers instead of going off on my own is the saving of fifteen more lives: twenty-five lives instead of ten.

Clearly, this is problematic. By hypothesis, the consequence of my joining the other rescuers instead of going off on my own is the overall saving of ten *fewer* lives: a hundred lives instead of a hundred and ten. Rejecting the Share-of-the-Total View, Parfit asserts the following principle, which he calls (C6):

(C6): An act benefits someone if its consequence is that someone is benefited more. An act harms someone if its consequence is that someone is harmed more. The act that benefits people most is the act whose consequence is that people are benefited most. (p. 69)

Parfit notes that (C6) revises our ordinary use of 'benefit' and 'harm' in several ways, one of which is exhibited in the following example. Ordinarily, if I save J's life, we would say that I benefit J; ordinarily we would say this even if someone else would have saved J's life in my absence. But according to (C6), in such a case I do not benefit J by saving his life, since J is not thereby benefited *more* (i.e. more than he would have been benefited in my absence).⁵ So (C6) sometimes invites us to describe the consequences of acts in counter-intuitive ways.

But Parfit argues that this peculiarity of (C6) is excused by the way in which (C6) helps agents to compare the consequences of their possible acts. As Parfit points out, if I am in a situation like the one just described, except that I have the additional option of saving K's *arm* instead of saving J's life (and no one else will save K's arm if I do not), then (C6) implies that my choice is between (1) doing, effectively, *nothing* (which is the consequence, according to (C6), of my saving J's life) and (2) saving K's arm. So according to (C6), the consequence of my saving K's arm instead of my saving J's life is the saving of an extra arm, with no loss of life. And this is a perfectly satisfactory result, since it exactly matches the implications of the situation as described. Parfit concludes that 'The First Mistake in moral mathematics is the Share-of-the-Total View. We should reject this view, and appeal instead to (C6)' (p. 70).⁶

⁵ According to R. M. Hare, in such a case not only do I not benefit J, but I do not even act at all, since 'To act is to make a difference to the course of events, and what the act is, is determined by what difference' (*Sorting Out Ethics*, Oxford, 1997, p. 164). I shall not evaluate this strong claim here.

⁶ Parfit's rejection of the Share-of-the-Total View echoes Donald Regan's rejection of the 'contributory consequences' approach in favour of the traditional 'marginal consequences' approach. Regan's discussion of these competing approaches (*Utilitarianism and Co-operation*, Oxford, 1980, pp. 13–17) is extremely illuminating.

III. THE SECOND 'MISTAKE'

The second 'mistake' is to consider an act in isolation from other acts with which it is connected by assuming, as Parfit puts it, the following:

(The Second Mistake) If some act is right or wrong *because of its effects*, the only relevant effects are the effects of this particular act. (p. 70)

Unfortunately, the way in which Parfit expresses this 'mistake' does not seem to be what's needed in order for it to be construed as a mistake – for what could the phrase '*its effects*' refer to, if not the effects of this particular act? As it stands, the second 'mistake' is evidently a tautology, and so Parfit must have in mind (as the second 'mistake') something other than the proposition quoted above.⁷ Therefore, let us follow Jackson⁸ in assuming that the 'its' is there erroneously, so that the second 'mistake' is actually the following:

(The Second Mistake) If some act is right or wrong *because of effects*, the only relevant effects are the effects of this particular act.

To see the upshot of making the second 'mistake', consider the following situation:

Case One. X and Y simultaneously shoot and kill me. Either shot, by itself, would have killed. (p. 70)

Parfit observes that 'According to (C6), neither X nor Y harms me', since neither makes me worse off than I would be if he were to act differently. (Of course, if *both* were to act differently, I would be better off because I would not die; but given the other's conduct, neither makes me worse off just by putting an extra bullet in me.) X does not harm me, and Y does not, either. If we make the second 'mistake', we conclude that neither X nor Y acts wrongly.

But Parfit calls this conclusion 'absurd', and adds that some see this as a *reductio* of (C6). But instead of rejecting (C6), Parfit proposes to supplement it with the following principle, which he calls (C7):

Even if an act harms no one, this act may be wrong because it is one of a *set* of acts that *together* harm other people. Similarly, even if some act benefits no one, it can be what someone ought to do, because it is one of a set of acts that together benefit other people. (p. 70)

Thus 'X and Y act wrongly because *they together* harm me. ... On any plausible theory, even if each of us harms no one, we can be acting

⁷ And other than what he says on p. 443, where he repeats verbatim the statement of the second 'mistake' already quoted.

⁸ Jackson, p. 52n3.

wrongly if we together harm other people' (p. 70).⁹ Not to realize this is to make the second 'mistake'.

IV. HOW PARFIT'S INDICTMENT OF THE SECOND 'MISTAKE' IMPLIES THAT PARTICIPATION MATTERS

This may already be evident, but it should prove worthwhile to pause here to make it explicit. Let us begin by reverting to a distinction casually introduced in section I, between making a difference and participating. Let us say that an agent *makes a difference* when her act benefits someone or harms someone, in the sense given to these terms by Parfit's principle (C6). And let us say that an agent *participates* when her act is one of a set of acts that together make a difference, regardless of whether her act itself makes a difference. It is clear that in Case One, X participates in killing me, and so does Y. But neither makes a difference. Now according to Parfit, to make the second 'mistake' is to think that because neither X nor Y makes a difference – i.e. because each merely participates – neither X nor Y acts wrongly. To avoid the second 'mistake', then, is to think that participation matters. Indeed it is precisely because X and Y *participate* in killing me that Parfit calls them both 'murderers' (p. 70).¹⁰

V. HOW PARFIT'S INDICTMENT OF THE FIRST 'MISTAKE' IMPLIES THAT PARTICIPATION DOES NOT MATTER

Recall that the first 'mistake' is the Share-of-the-Total View. We saw in section II that by crediting acts with producing benefits and harms they do not really produce – by crediting acts with making a difference when they really do not – the Share-of-the-Total View

⁹ Some support for this view can be drawn from American tort law. Courts have ruled that in over-determination cases, each agent is liable for all of the injury caused:

Where two or more causes combine to produce such a single result [i.e. an 'indivisible' one, such as death], incapable of division on any logical or reasonable basis, and each is a substantial factor in bringing about the harm, ... each of the causes is charged with responsibility for the entire harm. ...

... Such entire liability ... is imposed where either cause would have been sufficient in itself to bring about the result. (American Law Institute, *Restatement (Second) of the Law: Torts*, St Paul, Minn., 1965, ii, p. 440.)

From this it would follow that in Case One, X and Y are each liable for the entire harm of my death. To the extent that we construe liability as a proxy for culpability (only a rough parallel, admittedly), we may think that several people each act wrongly when they together harm other people.

¹⁰ The thesis that participation matters does not, of course, mean that participation always matters to such an extent as to be decisive (as it happens to be in Case One). It simply means that participation is a morally relevant factor to be considered alongside others (including, presumably, making a difference).

makes me think that in the First Rescue Mission, I have a moral reason to join the others instead of going off on my own. My claim in this subsection is that this is precisely *why* Parfit regards the Share-of-the-Total View as a mistake: because it says I have a moral reason to join the others – in other words, because it says my participation matters. It follows that Parfit's rejection of the Share-of-the-Total View – his characterization of it as the first 'mistake' – implies that participation does *not* matter.

Now it may be objected that I am reading too much into Parfit's rejection of the Share-of-the-Total View. Perhaps Parfit's indictment of the first 'mistake' does not rule out *every* view that implies that participation matters, but only those which, like the Share-of-the-Total View, ascribe to acts (such as my joining the others) the production of benefits and harms they do not really produce – for that is the type of mistake the first 'mistake' really is. This leaves Parfit free to accept some other view that implies that participation matters, as long as it does not make the kind of mistake that sinks the Share-of-the-Total View. For example, it leaves him free to appeal to a principle such as his own (C7) – in particular, the second sentence of it, which says that even if an act does not produce any benefit, it may be what an agent ought to do because it is one of a set of acts that together produce some benefit. Clearly (it might be thought) this way of saying that participation matters is untouched by Parfit's indictment of the first 'mistake' because it never credits an act with producing any benefit or harm it does not actually produce. Indeed it fully admits that certain acts do not produce any benefits or harms at all; but it maintains that there are morally significant characteristics of acts aside from the production of benefits and harms, and that one of these is *membership* in a set of acts that *together* produce benefits or harms.

This objection shows that one need not take the Share-of-the-Total View in order to maintain that participation matters. Instead, one can appeal to (C7), hoping to steer clear of Parfit's indictment of the first 'mistake'. And this approach may seem promising, for it may seem implausible what I claimed in the first paragraph of this section: that Parfit regards the Share-of-the-Total View as a mistake because it says that I have a moral reason to join the others. After all, this objection reminds us that Parfit's own principle (C7) implies that I have such a reason; and surely (it might be thought) Parfit's indictment of the first 'mistake' – even if it has some implications beyond those that Parfit makes explicit – does not rule out *this* principle, which he asserts just a few pages later.

So this objection claims that Parfit's indictment of the first 'mistake' does not extend so far as to implicate views such as Parfit's principle (C7). But a moment's reflection shows that problems lurk in the appeal

to (C7). To see this, consider an appeal to (C7) in the context of the First Rescue Mission. We have seen that if I join the others and defend my decision by pointing out that I thereby save twenty-five lives, then I make a mistake. Now suppose that I join the others and defend my decision in the way suggested in the objection given above, by appealing to (C7). It seems clear that if the Share-of-the-Total View misleads an agent in the First Rescue Mission, then surely (C7) does, too.¹¹

Now our objector may offer the following reply. The First Rescue Mission shows the Share-of-the-Total View to be a mistake in just the way that Parfit says it's a mistake – it credits acts with producing effects they do not really produce. (C7) does no such thing, in the First Rescue Mission or anywhere else. (C7) emerges unscathed, because the implications of Parfit's indictment of the first 'mistake' are limited by the narrowness of the error embodied in the Share-of-the-Total View.

However sensible this reply may initially appear, I believe that it rests on a failure to see what's really wrong with the Share-of-the-Total View. Notice that if I take the Share-of-the-Total View in the First Rescue Mission, one problem that does *not* result is that I then misunderstand the brute facts of the situation. Taking the Share-of-the-Total View does *not* make me think that there exists some causal chain, previously unnoticed by me and possibly undiscoverable by me, beginning with my joining the others and terminating in the saving of twenty-five lives; it does not saddle me with any such mystical (or, at least metaphysically suspicious) beliefs at all. All it makes me think is that my joining the others has, as an act, a moral worth comparable to that of an act of actually saving twenty-five lives. In short, it leads me to think that I have a moral reason to join the others. And this is the *only* thing that it makes sense to fault the Share-of-the-Total View for doing. So if we regard the Share-of-the-Total View as a mistake, then

¹¹ It might be thought that (C7) does not necessarily mislead an agent in the First Rescue Mission, once the 'set of acts' to which (C7) refers is properly characterized. For if we take one 'set of acts' to involve my joining the others, and we take another 'set of acts' to involve the other four rescuers' standing on the platform and my going off on my own, then (C7) does not say that I have a moral reason to join the others rather than going off on my own, and thus does not mislead me. But if (C7) is to be interpreted in this way – in such a way that (C7) does not say that I have a moral reason to join the others rather than going off on my own – then it cannot be adduced as an example of a view that implies that participation matters, and (C7) ought not to be mentioned in connection with the objection described and answered in this subsection. But this would not change the content of the objection itself, or undermine my response to it. For if I join the others and defend my decision by appealing to some (C7)-like principle that *does* imply that participation matters, then clearly I have been led astray, in the sense that I have been led to make the wrong decision – indeed no less than if I had made the cruder error of thinking that, in joining the others, I would be causing fifteen more lives to be saved than if I were to go off on my own. I am indebted to an anonymous referee for this journal for making clear the need for this note.

it cannot be on the basis of some disputed matter of fact as to how many lives my joining the others *really* saves; it must be because the Share-of-the-Total View tells me that I have some moral reason to join the others. It follows that the fallout from Parfit's indictment of the first 'mistake' cannot be restricted in the way suggested by our objector. On the contrary, Parfit's indictment of the first 'mistake' must imply the rejection of (C7), simply because it says that I have a moral reason to join the others.

At this point, it may be replied that when we regard the Share-of-the-Total View as a mistake, then we really *are* concerned about a matter of fact, not the more distant question of what moral reasons I do and do not have. After all, Parfit says that he's concerned to correct some mistakes in 'moral mathematics', and he insists that the number of lives I save by joining the others is not indeterminate, but zero. Surely this, our objector might say, implies that he *is* talking about some matter of fact. But Parfit's insistence that the answer is zero actually supports the hypothesis that what is really wrong with the Share-of-the-Total View is that it makes me think that I have a moral reason to join the others – for zero is the only answer that is guaranteed to make me think that I have *no* moral reason to join the others; any answer other than zero leads me to think that I have some moral reason, however small, to join the others.

Let us look at the matter from a slightly different angle: What is the *best explanation* for Parfit's indictment of the first 'mistake'? One answer, my answer (which is my main point in this section), is that Parfit wants to draw attention to a certain *moral* error having to do with what moral reasons agents have to perform certain acts. Another answer, the one given by the objection now under consideration, is that Parfit wants to draw attention to a certain *non-moral* factual error having to do with what consequences result from certain acts. Now it seems likely that the error I focus on (the one having to do with what moral reasons agents have to perform certain acts) is far more of a problem in moral thinking than is the other error (the one about what consequences result from certain acts). What makes the most sense is to conclude that Parfit's indictment of the first 'mistake' applies not only to the latter error, which hardly anyone would fall for, but also to the former one, which is genuinely tempting. Otherwise, we effectively consign Parfit's indictment of the first 'mistake' to being oddly tangential, or even irrelevant, to the way in which ordinary people think about the moral significance of the consequences of their acts.

Finally, it may be objected that I have overlooked another, more plausible, explanation for Parfit's indictment of the first 'mistake'. The last sentence of (C6) says that '[t]he act that benefits people most is the act whose consequence is that people are benefited most', and so (C6)

implies that I have a moral reason to go off on my own, and hence a moral reason not to join the others. With this in mind, it might be claimed that Parfit's reason for criticizing me for taking the Share-of-the-Total View and joining the others in the First Rescue Mission is not that I thereby embrace the view that participation matters, but that I thereby neglect the just-mentioned implication of (C6).

Now I concede that this reading of Parfit's text is a reasonable alternative to the one I have been relying on so far. But I maintain that this alternative reading of Parfit's text reinforces, rather than undermines, my claim that Parfit's rejection of the Share-of-the-Total View implies that participation does not matter. Once (C6) is read in the strong way that this reading of Parfit's text requires – namely, as supplying conclusive moral reasons for and against performing certain acts – then moral reasons that come from the prospect of making a difference end up trumping those that come from the prospect of participation. In effect, there is simply no room left for the view that participation matters.

In reply, it might be thought that I have read too much into (C6); and that this principle, instead of supplying *conclusive* moral reasons for and against performing certain acts, supplies only *pro tanto* reasons for and against performing certain acts.¹² And if this is how (C6) is to be read, then there is plenty of room left for the view that participation matters, since presumably in some cases the *pro tanto* reasons supplied by (C6) could be overridden by participation-based reasons. On this reading of (C6), Parfit's reason for rejecting the Share-of-the-Total View is not that participation does not matter, but rather that the Share-of-the-Total View makes participation appear to matter more than it really does, by making participation-based reasons outweigh (C6)-based reasons in a case in which they should not.

But this interpretation of (C6) does not sit well with the rest of Parfit's discussion. For if Parfit regards the reasons supplied by (C6) as capable of being overridden by participation-based reasons, then there must be a variant of the First Rescue Mission in which the benefit to be gained by my going off on my own is positive but so small that Parfit would say that I ought to forgo that benefit and just join the others, despite thereby not making any difference at all. But when Parfit criticizes me for taking the Share-of-the-Total View in the (unmodified) First Rescue Mission, he does not in any way imply that if the benefit to be gained by my going off on my own were smaller, then I would be right to forgo it and join the others. Indeed the thrust of his reasoning is that the mere fact of *some* benefit to be gained by my going

¹² This interpretation of (C6) was suggested to me by an anonymous referee for this journal.

off on my own is enough to make it the case that I ought to do that instead of joining the others. Moreover, as indicated above, when he considers my joining the others, he says that the number of lives I save is zero; and he does *not* say anything like this: that although the number of lives I save is zero, there is still something right about the Share-of-the-Total View's implication that the prospect of my (ineffectual) participation counts, however lightly, in favour of my joining the others. So it is hard to see how Parfit might regard (C6) as supplying only *pro tanto* reasons, capable of being overridden by participation-based reasons.

It follows that we must construe (C6) in the strong way that I proposed at the beginning of this subsection. But if we do so, then we also thereby construe it so that it is strong enough to imply that participation does not matter. Thus, whether we explain Parfit's indictment of the first 'mistake' as I originally proposed earlier in this section or by appealing to (C6), we find that his indictment of the first 'mistake' implies the rejection of any view that implies that I have a moral reason to join the others. So his indictment of the first 'mistake' implies that participation does *not* matter, while his indictment of the second implies that it does.

VI. THE 'MISTAKES' REVISITED

In retrospect, some hint of this conflict can be seen even in Parfit's discussion of Case One, in which X and Y simultaneously shoot and kill me. As we saw at the beginning of section III, Parfit acknowledges the relevance of his principle (C6), which includes the proposition: 'An act harms someone if its consequence is that someone is harmed *more*' (my emphasis). Since neither X nor Y harms me more than I would be harmed if he were to act differently, (C6) implies that neither X nor Y harms me. So far, so good. But from this observation, it is tempting to infer that neither X nor Y acts wrongly, which is to make the second 'mistake'. (Indeed, Parfit introduces Case One precisely in order to characterize the second 'mistake'.) This connection between (C6) and the second 'mistake' would be unremarkable were not (C6) presented by Parfit as the remedy for the first 'mistake'.

Nor would this awkwardness be avoided if the order of exposition were reversed – for suppose Parfit were to begin with Case One and the 'mistake' that goes with it (the second 'mistake'), and were then to present (C7) as the remedy for that 'mistake'. Then in the subsequent discussion of the First Rescue Mission, he would have to acknowledge the relevance the second sentence of (C7), which says that even if an act does not produce any benefit, it may be what an agent ought to do because it is one of a set of acts that together produce some benefit.

And from this it would be tempting to infer that in the First Rescue Mission, I have some moral reason to join the others – which is to make the other ‘mistake’ (the first ‘mistake’). (Indeed were Parfit to proceed in this way, he would not even need to introduce the Share-of-the-Total View in order to characterize the first ‘mistake’; he could simply appeal to (C7), just as, in fact, he simply appeals to (C6) in order to characterize the second ‘mistake’.) The circle is complete: for each of the two mistakes, its remedy is the other’s root.

VII. A POSSIBLE RESOLUTION EXAMINED

There is one other passage in Parfit’s discussion of ‘moral mathematics’ that we should consider, because it may seem to offer a resolution of this inconsistency. In his principle (C13), Parfit says that when a person’s inclusion in a benefit-producing group does not increase the benefits produced by that group, then ‘he has no moral reason to join this group’ (p. 83). This is consistent, of course, with our understanding of Parfit’s indictment of the first ‘mistake’. But (C13) mentions only benefits, not harms; and Parfit’s indictment of the first ‘mistake’, being driven by the First Rescue Mission, is also concerned only with benefits, not with harms. Meanwhile, Parfit’s indictment of the second ‘mistake’ is driven by Case One, in which there are harms but not benefits. Can it be Parfit’s position that participation matters in the production of harms but not in the production of benefits?

This seems unlikely, for a couple of reasons. First, there is textual evidence that Parfit does not occupy this position, since (C7) implies that participation matters not only in the production of harms, but also in the production of benefits. (It says that ‘even if some act benefits no one, it can be what someone ought to do, because it is one of a set of acts that together benefit other people’ (p. 70).) So even if we understand Parfit’s indictment of the first ‘mistake’ to be concerned only with benefits, as (C13) is, then Parfit still mires himself in a contradiction about participation, since he says both that participation in the production of benefits matters (in (C7)) and that it does not (in (C13)).

Second, the position that participation matters in the production of harms but not in the production of benefits presupposes a curious asymmetry between benefits and harms. If both are morally significant, as far as consequences are concerned, then why is not participation in the production of the one as morally significant as participation in the production of the other? We saw that Parfit’s principle (C13) says that when a person’s inclusion in a benefit-producing group does not increase the benefits produced by this group, then ‘he has no moral reason to join this group’ (p. 83). Why does it not follow from this that

when a person's inclusion in a *harm*-producing group does not increase the *harms* produced by that group, then he has no moral reason *not* to join that group? (It would not follow that he has some moral reason *to* join, but that is not the point.) Now perhaps Parfit would resolve the inconsistency in this way – by amending (C7) and claiming an asymmetry between benefits and harms. But this position is so odd that it would be a stretch to ascribe it to Parfit without some affirmative evidence that it is the resolution he would choose.

VIII. CONCLUSION: THE WIDER RELEVANCE OF PARFIT'S ERROR

In closing, I want to review the main claims of this paper and to reflect on the relevance of them to consequentialist thought in general. After summarizing Parfit's discussions of the first and second 'mistakes' in moral mathematics, I showed that Parfit's discussions respectively imply that, on the one hand, participation does not matter and that, on the other, it does. I examined one possible resolution of this conflict, but concluded that it is unsatisfactory.

The inconsistency exposed in this paper concerns not only Parfit's theory of moral mathematics, but also consequentialist thought in general, for Parfit's inconsistency is not only a significant (if regrettable) factor in his explication of consequentialism, but also an instance of a deep tension to which much contemporary discussion of consequentialism is addressed: the tension between so-called 'act-' and 'rule-' versions of consequentialism. This issue is complicated enough for any brief characterization of it to be immediately (and deservedly) suspect, but for our purposes it is sufficient to regard act-consequentialism as holding that the moral assessment of an act depends primarily on its consequences, and rule-consequentialism as holding that the moral assessment of an act depends primarily on the consequences that would result if a rule allowing it were generally taught and/or accepted.

To see how Parfit's inconsistency relates to this conflict, notice that Parfit's indictment of the second 'mistake' has rule-consequentialist undertones, since his principal example of the second 'mistake' is the judgement that in Case One, neither of the shooters acts wrongly. While this judgement is consistent with act-consequentialism, it may be contested on rule-consequentialist grounds. Specifically, consider the following rule: It is not wrong to participate in the harming of others, as long as one's own involvement happens to be superfluous, given the involvement of others. Now if this rule were generally taught and/or accepted, then the consequences would surely be quite bad, since there would be no prohibition on behaviour such as that of X and

Y in Case One. So although Parfit does not marshal explicitly rule-consequentialist arguments in order to justify his indictment of the second ‘mistake’, rule-consequentialist grounds provide a perspective from which to understand his position in a way that act-consequentialist grounds do not.¹³

On the other hand, Parfit’s discussion of the first ‘mistake’ is more congenial to act-consequentialism. Consider, for example, the weight Parfit puts on the fact that the consequence of my helping the other four rescuers is worse than the consequence of my going off on my own – this observation, which forms the nucleus of Parfit’s discussion, invokes no rule-consequentialist considerations. Moreover, Parfit’s affirmation of his principle (C6) as a remedy for the first ‘mistake’ confirms his commitment to a straightforward comparison of the available acts’ individual consequences, again with no reference to rules, tendencies, generalities, or any other hallmarks of rule-consequentialism. Now it must be admitted that Parfit’s *judgement* in regard to the First Rescue Mission (i.e. that I ought to go off on my own) is not distinctly act-consequentialist; a rule-consequentialist could certainly offer the same judgement. But Parfit’s *reasoning* – especially the aspects just mentioned – invokes the straightforward comparisons that are characteristic of act-consequentialism and bears no signs of rule-consequentialist influence.

We may conclude, then, that the inconsistency in which Parfit’s indictments of the first and second ‘mistakes’ ensnare him can be understood as an instance of the tension in contemporary consequentialist thought between ‘act-’ and ‘rule-’ versions of consequentialism. It is as if Parfit were moved by act-consequentialism in his discussion

¹³ We saw in note 9 that American tort law effectively implies that participation matters. It is also instructive to note the influence of rule-consequentialist considerations on this area of the law. In 1948, a court faced a case in which two hunters shot at a quail flushed by a third hunter and hit that third hunter with birdshot. The court held each of the shooters liable for all of the injury (even though possibly only one of them actually shot the plaintiff), saying that once the plaintiff had proved that the defendants had jointly injured him, it was up to each defendant to prove that he himself had not injured the plaintiff. Here is how the court justified shifting the burden of proof in this way (though the emphasis is added):

When we consider the relative positions of the parties and *the results that would flow* if plaintiff was required to pin the injury on one of the defendants only, a requirement that the burden of proof on that subject be shifted to defendants becomes manifest. (*Summers v. Tice*, 33 Cal. 2d at 86, 199 P. 2d at 4)

This support that American tort law derives from rule-consequentialism, combined with the support that American tort law provides for the view that participation matters, tends to confirm the suggestion in the text that rule consequentialism is especially congenial to the view that participation matters. For an illuminating discussion of the issues raised by *Summers v. Tice*, see Judith Jarvis Thomson, ‘Remarks on Causation and Liability’, *Philosophy & Public Affairs*, xiii (1984). She briefly mentions the possible role of rule-utilitarian considerations in apportioning costs at 114.

of the First Rescue Mission but – faced in Case One with some of its less attractive implications – then recoiled from it and took up the perspective of rule-consequentialism instead. Seen in this light, Parfit's inconsistency is eminently explicable – for act-consequentialism and rule-consequentialism each have some strong appeal; otherwise, they would not each remain the subject of considerable attention and development. But to draw on each of them, however subtly, is ultimately untenable. It is tempting, but futile, to try to have it both ways.¹⁴

eggleston@ku.edu

¹⁴ My thoughts on *Reasons and Persons* have been greatly influenced by conversations I have had with David Gauthier, and I want to thank him, Dale Miller, Donald Bruckner, Lad Sessions, Alastair Norcross, and an anonymous referee for this journal for giving me extensive comments on earlier drafts of this paper. I should also mention that I may well not have thought of writing the present paper if I had not first read the paper by Frank Jackson cited above. Finally, I want to thank the audience in front of whom I presented this paper on December 30, 1999, at a session of the Eastern Division meeting of the American Philosophical Association.