

# Sets of acts<sup>\*</sup>

Martin E. Sandbu  
Department of Legal Studies and Business Ethics  
Wharton School of Business  
sandbu@post.harvard.edu

26<sup>th</sup> August, 2007

Paper presented at the annual meeting of the American Political Science Association,  
Chicago, IL, August 2007

## 1. Introduction

What we do together matters morally. Perhaps the greatest moral challenges facing people in our part of the world involve the effects of what we do together or fail to do together. This is because the phenomena that will have the greatest effect, and potentially the greatest negative effect, on human lives in the foreseeable future are those which are either caused by what we do together or can only be forestalled by what we do together. I have in mind the processes of climate change, disease pandemics and epidemics, and economic development. The actions of each individual make little or no difference to these phenomena, but the actions of all do.

Our standard theories of morality and rationality struggle with these situations. The main problem is that individual actions are ineffectual even when collective action is not, which drives a wedge between the individual perspective of most theories both of morality and of rational choice, and the collective level at which it is possible to act effectively. In this essay I offer some observations prompted by Derek Parfit's suggestion that the rightness or wrongness of an act can depend on the rightness or wrongness of the set of acts of which it is part. This suggestion raises two sets of question. The first is when we can say that a *set* of acts is wrong, apart from the trivial case where it is wrong because it is a set of individually wrongful acts. The second is what the fact that a set of acts is wrong, when it can be established, entails for the individual acts that constitute the set. These questions are more general than the question of collective action. They can equally well be asked of sets of acts performed by a single individual over time. We have more solid answers to these questions when asked in that context. The same answers, I will argue, apply to sets of acts performed by several individuals.

---

<sup>\*</sup> I am grateful to Annie Stilz and Micah Schwartzman for very useful feedback on a draft version of this paper.

## 2. Sets of acts

Derek Parfit makes what I shall call a *consequentialist set-of-acts claim*:

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 people.<sup>1</sup>

The set-of-acts claim serves to account for our intuitions in cases where each *individual* action does not perceptibly contribute to good outcomes, but where all the actions taken together do so contribute. One way this may be true is in the case of imperceptible individual effects. Suppose there are a large number of wounded people in the desert, and a large number of people, each with a pint of water and with one water cart in common. We can each pour a pint into the water cart, which can be driven out to the desert, and the water will be evenly distributed. Each extra pint would only add an imperceptible amount of water for each wounded person, but if enough of us contribute our water, we will significantly reduce their suffering. Even if the effects of a single pint are imperceptible, each person should contribute their pint.<sup>2</sup>

The importance of sets of acts need not be confined to consequentialism. Non-consequentialists may think:

Even if an act is not wrong by itself, this act may be wrong because it is one of a set of acts that together are wrong. Similarly, even if some act is not what someone ought to do in isolation, it can be what someone ought to do, because it is one of a set of acts that together ought to be done.

We may call this the generalized set-of-acts claim. We shall discuss the consequentialist version for simplicity; similar points can be made for the generalized version. Either version of the claim is underdefined. The consequentialist set-of-acts claim does not say that an act that is one of a set of acts that harm other people *is* wrong because of that fact, but that it *may* be wrong because of that fact. We should ask when it is wrong because of that fact. That fact does not always make it wrong. For it could be that an act is one of a set of acts that together harm other people without it being the case that someone ought not to do that act, as in:

*The stock market crash.* I have my savings invested in the stock market of a poor country. Should I lose a large part of these savings, my family and I will suffer. One day there is a *self-fulfilling panic* in the stock market: People think the stock prices will go down, therefore they sell, driving the prices even further down. It becomes clear to me that within a few hours, my holdings could fall to a fraction of their original value. Such a fall in stock market prices would gravely harm the inhabitants of the poor country. Many would lose their jobs or their savings, and some would starve. I sell my stocks as soon as I can.

---

<sup>1</sup> Parfit, *Reasons and Persons*, 1984, p. 70.

<sup>2</sup> Parfit, 76, citing Jonathan Glover.

In this case, my act is one of a set of acts that together are seriously harming other people. Am I acting wrongly? I am not. If I hold on to my stocks, no good will come of it. But my family and I will suffer. I am not morally required to make a useless sacrifice.

Parfit recognises this. He proposes to appeal to the following principle:

When (1) the members of some group would make the outcome better if *enough* of them act in some way, and (2) they would make the outcome *best* if *all* of them act in this way, and (3) each of them both knows these facts and believes that enough of them will act in this way, then (4) each of them ought to act in this way.<sup>3</sup>

In the Stock Market Crash, this principle rightly does not say that I ought to hold on to my stocks if everybody else is selling. But it does say that I ought to stop selling if I believe that enough people stop selling.

Even if this principle is true, it is inadequate. Consider Parfit's case of the *Harmless Torturers*:

*The Harmless Torturers.* A thousand torturers have a thousand victims. At the start of each day, each of the victims is already feeling mild pain. Each of the torturers presses a button, thereby turning a switch once on each of a thousand instruments. Each turning of a switch affects some victim's pain in a way that is imperceptible. But, after a switch has been turned a thousand times, it has inflicted severe pain on the victim.<sup>4</sup>

Here (1) and (2) hold. The outcome would be better if enough of the torturers refrain from pressing the button, and it would be best if all of them refrained. Suppose now that (3) does not hold:

*The Skeptical Torturer.* All the torturers know the facts of *The Harmless Torturers*. Some torturer believes that not enough of them will refrain from pressing their button.

Ought he to refrain, according to the principle? We must think that either (a) this principle is silent on whether this torturer ought to refrain; or (b) "when" means "when and only when" so the principle entails that it is *not* the case that the torturer ought to refrain. This latter conclusion is absurd. On interpretation (b) the principle must be rejected. So we should choose interpretation (a). On interpretation (a) the principle is inadequate by itself because it is incomplete. We need a more general principle.<sup>5</sup>

---

<sup>3</sup> Parfit, p. 77f.

<sup>4</sup> Parfit, p. 80f.

<sup>5</sup> Parfit (81) offers the following principle, which would account for the wrongness of the skeptical torturer: When (1) the outcome would be worse if people suffer more, and (2) each of the members of some group could act in a certain way, and (3) they would cause other people to suffer if *enough* of them act in this way, and (4) they would cause these people to suffer *most* if they *all* act in this way, and (5) each of them both knows these facts and believes that enough of them will act in this way, then (6) each would be acting wrongly if he acts in this way.

While this principle has the correct implication that the the single torturer is acting wrongly, it has the incorrect implication that each person who sells their stocks in the Stock Market Crash is acting wrongly.

McMahon offers a different principle that has very similar implications. He offers a prudential principle for when one has reason to do the things that together with the actions of others would make the outcome better. He calls this the “Principle of Collective Rationality”:

One has sufficient reason to contribute as provided to a cooperative scheme that produces something one regards as good if the value to one of the outcome of the scheme, when one’s contribution is added to the others that will actually be made, exceeds the value to one of the noncooperative outcome. Conversely, one has sufficient reason to defect from a cooperative scheme if the value to one of the noncooperative outcome exceeds the value to one of the outcome of the scheme (when one’s contribution is added to the others that will actually be made).<sup>6</sup>

This is not a moral principle. But it is structurally equivalent to Parfit’s claim. It applies to Prisoner’s Dilemmas, Assurance Games, and Co-ordination games, and claims that in these games an individual has a *prudential* reason (not moral) to do the action that together with the appropriate actions of others would make the outcome better, judged from that individual’s values (not an objective moral judgment). Individuals who value outcomes that are morally better in Parfit’s sense would have sufficient reason, on McMahon’s principle, to do exactly the same as what they ought to do, on Parfit’s principle (except that McMahon does not require the outcome to be the *best* when *all* cooperate). McMahon’s principle, in particular, does not give the skeptical torturer a sufficient reason to refrain from pressing his button, even if he values an outcome with less pain as better than one with more pain. McMahon’s principle can also be interpreted as either incomplete or exclusive. Since the exclusive interpretation leads to the absurd conclusion that the skeptical torturer *does not have* a reason to refrain from pressing his button, we must read it as incomplete.

We need a principle that gives the right answers in cases where not everyone believes that enough will perform the actions that together will make the outcome better (or that together will be right, if we are non-consequentialists). These are cases that are difficult both practically and theoretically. They are difficult practically because individual decision-makers have reasons to act in ways that together make the outcome worse. Here we are concerned with the theoretical difficulty. The theoretical difficulty can be informally stated in the following way. The actions of all taken together are make the outcome worse. Yet the actions of each make the outcome the best, given everyone else’s actions. This is a theoretically important challenge even in cases where the values that are thus frustrated are immoral, such as when a group of criminals fail to collaborate to carry out a crime. This is a defeat of rationality even if it is a victory for morality. The theoretical challenge is all the more important when the outcomes we fail to achieve are outcomes that are morally valuable. In those cases the challenge to rationality is also a moral challenge.

---

<sup>6</sup> McMahon, *Collective Rationality and Collective Reasoning*, 2001, p. 21f.

### 3. Sets of actions over time

There is something puzzling about the set-of-acts claim. For what if we ask: “Why should it matter for the evaluation of an act that it is part of a set of acts?” Parfit asserts that it is a mistake to think that “[if] some act is right or wrong *because of its effects*, the only relevant effects are the effects of this particular act.” But why is this a mistake? Similarly, why is it irrational, on McMahon’s view not to cooperate in prisoner’s dilemma’s or assurance games? So long as my act makes the outcome better, why is it relevant that it is part of a set of acts that makes the outcome worse? Why is it more relevant that the set of acts of which it is part makes the outcome worse than it is that an action of *yours*, unrelated to mine, makes the outcome worse? To answer these questions properly, we must not only show when the set-of-acts claim is plausible, we must also show why it is plausible in those cases and not in others.

Discussions of sets of act usually focus on sets of actions performed by different people at the same time. But the force of the set-of-acts claim is best seen when applied to sets of actions performed by the same person at different times. Consider:

*Watering plants.* John promises Sam to care for Sam’s plant during Sam’s two-month vacation. The plant will survive if John waters it on enough days, and it will be best cared for if he waters it every day. If he waters it more than a little every day, it will die from too much water, but it can survive without water for several days at a time without any noticeable difference. Every day, John can walk over to Sam’s and water the plant. Given his behaviour on all other days, his behaviour on any given day does not affect the outcome. It takes a considerable amount of time out of his day to go over to Sam’s.

John has an obligation to water the plant. But suppose he disputed this. He could give as an argument Parfit’s principle, *applied to his successive selves over time*. Label John on day 1 John-1, John on day 2 John-2, and so on. John (at all times) accepts Parfit’s principle applied to himself over time:

When (1) John-1, John-2... would make the outcome better if *enough* of them act in some way, and (2) they would make the outcome *best* if *all* of them act in this way, and (3) each of them both knows these facts and believes that enough of them will act in this way, then (4) each of them ought to act in this way.

More generally, this means:

When (1) a person would make the outcome better if *on enough days* he acted in some way, and (2) he would make the outcome *best* if he acted in this way *every day*, and (3) every day, he knows these facts and believes that on enough days he acts in this way, then (4) on every day, he ought to act in this way. Suppose John thinks on day 1 that he will not act in this way on enough days. Perhaps he has a realistic view of his own laziness. Then this principle does not tell him to act in this way on day 1. On day 2, he may still think that he will not act in this way on enough days. And he has even better reason to think so on day 2, since he knows that on one day when he could have acted this way, he did not. And so on. John may even understand this iterative argument, and use that as his reason why on day 1 he does not think he will act this way

on enough days. So Parfit's principle, applied to successive selves, does not give any of the John-*t*'s a moral reason to water the plant, given what all the other John-*t*'s are doing.

John is mistaken. John does have an obligation to care for the plant. This is why John-1, John-2, *etc.* also have an obligation to water it on their respective days. But why is John mistaken?

Consider:

*Memento.* The next time Sam goes on vacation, he asks Leonard to care for his new plant. Leonard suffers from severe anteretrograde amnesia. This means that he has no capacity to encode short-term memories so that they are available to his recollection later. Every day is for him a new day. He remembers what happened to him up until the day he suffered the stroke that caused the amnesia, and he knows his condition. But he does not remember what he did or what happened to him yesterday. Like John, he promises to care for the plant. He does not water it.

Like John, Leonard justifies his actions every day with the same argument as John. Leonard-2 does not know what Leonard-1 has done. Nor can he commit Leonard-3 *et cetera*. As a result, on no day does Leonard believe that he will water the plant on enough days. If he waters it today, it will not make a difference, so the consequences of the single action do not mean that he must do it. And given his beliefs, the consequences of the *set* of his acts are not reason that he ought to water the plant.

Is Leonard more successful than John in justifying his failure to do the set of acts that makes the outcome better? I believe he is. The difference between Leonard and John is that we think John capable of planning and carrying out a sequence of actions over time in a way that Leonard is not. In both cases, we may admit that John-*today* and Leonard-*today* are not acting wrongly, given what the other John-*t* and Leonard-*t* are expected to do. But John, seen as a single moral subject, is acting wrongly. We attribute to him as a single moral subject the whole sequence of actions which *together* are wrong. The reason why the set of acts matters is that it is John's set of acts, and that John ought not to perform that set of acts. Each individual action is wrong *because* it is one of the set of acts that together constitute wrongdoing by John.

In the case of Leonard, there is less reason to treat the Leonard-*t*'s as one moral subject to whom we can attribute the whole sequence of actions. The relationships between Leonard-1, Leonard-2 *etc.* are too loose. There seems not to be a single moral subject Leonard who is performing the whole sequence of acts. Because there is no one who can be said to be performing the whole sequence of acts, there is no one who can be said to be acting wrongly in performing it. So each individual act is not one of a set of acts that together constitute wrongdoing by someone. If we think Leonard is less culpable than John, I believe this must be the reason why we think so.

We may say more about why John is properly seen as a single moral subject. We should not think that John's argument has no merit. He can rightly say: "You claim that there is a single agent named John. But there is no single agent, no decision-maker who can make all the decisions you attribute to John. I-here-and-now (John-today) cannot decide what 'my' previous selves did. Nor can I decide what 'my' future selves will do. I can try to

influence them, to be sure, but I cannot make the act done. So you are asking me to believe in a mysterious entity separate from the only existing decision-makers there are.”

The answer we could give to John-today is that he does in fact recognise a single moral agent John. We see him as a single moral subject insofar as we think it is in fact in his moral power to both commit his future selves, and “take responsibility” for his former selves’ failings, as we might say, perhaps by trying to make up for them. We know that he has these moral powers because he exercises them normally to serve his own interests and what he sees as his long-term goals. We can point to how he behaves in other parts of his life, where he undoubtedly pursues projects over long periods of time, makes long-term plans, and so on, and thinks of these projects and plans as *his* plans. He thinks of himself every day, that is, as *one* agent of which every moment’s decision-maker is the present representative.

This explains why John-*t*, on every day *t*, ought to water Sam’s plant. I believe it is also why he has *more* reason, not less, to do so if he has not done so in the past (provided it is not too late to make a difference if he does it every day from now on). The reason why each John-*t* has certain obligations is that the single agent who is John has an obligation to care for the plant, and that involves a set of actions by him at different times. And the reason why it is correct to attribute obligations to such a single agent is that John everyday has the requisite intentions and attitudes for there to be such an agent: His thoughts—at least enough of them—are in terms of “me through time,” not in terms of “me-here-and-now.”

Consider a third case:

*Living in the present.* Michael is a very light-hearted man, a hedonist with inherited wealth. He lives in the present. He does not care about the past, and if he you ask him to make commitments about the future, he may do so to humour you, but they mean nothing to him. These features about him make him very whimsical and inconsistent over time. When you point this out to him, he is indifferent. He may say “I was a different person then, and I’ll be a different person in the future.” He is not selfish, in fact he cares as much about others as about himself, and has been known to give away large amounts of money to needing strangers when he thinks that will do them good. He does not water the plant.

Every day, Michael fails to water the plant, because every day he thinks he will not do so on enough days for it survive, and so whether he does it today or not will not make any difference to the outcome. We may point out to him that if he watered it every day, his act today would be part of a set of acts that *would* make the outcome better. He responds with the same argument as John and Leonard.

Michael cannot justify his failure to act with this argument. The reason is not that he sees his successive selves as part of a single moral agent. He does not. We hold him responsible for the whole sequence of actions because he *could* have integrated his successive selves. He has simply not done so. He has not exercised a moral power that we do attribute to him.

But if these are the conditions for a set of acts mattering morally even when each act does not, then the same conditions can explain why a set of acts by different people can matter.

It can matter when we can attribute the set to a single collective agent—a plural subject. For again, the presence of a plural subject is just the existence of the appropriate attitudes in different minds. In the next section I discuss what these attitudes are.

#### 4. Collective action

The intuition that lies behind both Parfit's and McMahon's principles is the following. There are many situations in which I as an individual can do nothing to make an outcome better. But in some of these situations, *we together* can make the outcome better by together doing certain things. It is natural to say in some of these situations that *we* fail to do what *we* ought to do if *we* do not together do those things. And it is natural to say this even as we recognize that there is no single individual of whom it is true that he fails to do what he ought to do. That is, all may be acting wrongly together, while each is acting rightly individually. We often make judgments such as these. They are paradoxical. If no one is doing something he ought not to do, or is failing to do what he ought to do, what is left? What is the "we" that has duties that are not the duties of any individuals?

In the previous sections I showed how the same judgments are not paradoxical in the case of individual actions over time. John may today be acting rightly given what John has done in the past and will be doing on future days. But this behaviour is part of John's acting wrongly over time. The same is true of Karl and Michael. There is nothing mysterious in this, if we accept the idea of John (over time) as the subject of all the acts. John's obligations are not simply the aggregation of the individual obligations of John-*t* at each time *t*. They are obligations *of John*, and can be logically prior to the obligations of each of John's successive selves. Indeed the typical case is that the obligation of John's successive selves derive from the obligations of John-over-time, rather than the other way round.

On our common sense, pre-theoretical view, the same is true for groups of people. There can be obligations of groups of people that are logically prior to the obligations of individuals. "We" can *together* have obligations that are not aggregations of logically independent individual obligations of each of us. I believe that the common sense view is largely correct. A set of acts is wrong when that moral subject who unifies the acts is acting wrongly in performing the set of acts. This is true of John and Michael, but not of Leonard. The same analysis can be applied generally to sets of acts, whether by the same person over time, or by different people. The claim is this. An act can be wrong if it is one of a set of acts that together are wrong, even if the act by itself is not wrong. A set of acts can together be wrong if the set of acts can be attributed to a unified subject. When unified in this way, the set of acts can be wrong even if none of the individual acts is wrong by itself. And the wrongness of the set of acts is independent of whether the subject performing it is constituted of different individuals or of the same individuals' successive selves.

Thus for the set-of-acts claim to apply to acts by different people, we have to say that a group of people are together acting wrongly in a sense that does not derive from the individuals acting wrongly by themselves. More generally, we have to say that a group can have obligations that are not just a function of the individual obligations of its members. In the next section we discuss how this could be. Here we should note one

logical implication. It is that it is conceptually perfectly possible that there could be a *collective* moral failure even if no *individual* fails to do what he or she ought to do, given what everyone else is doing. The moral failure that the set-of-acts claim tries to capture could be interpreted in at least two ways:

- (1) There is a group of people who constitute a “we”: They share certain intentions that are the intentions of *all* to do certain things *together*. This makes *the group* capable of performing shared actions. This capability generates a duty for the group to perform shared actions that are morally required of *it* as a moral subject. The moral failure is the failure of the group to perform such shared actions that are morally required. It is analogous to John’s wrong in section 3.
- (2) There is no group of people who constitute a “we” in the sense of sharing intentions to certain things together. But there are some individuals who, if they did constitute a “we” in this sense, would as a group have the capability of performing certain shared actions that would be morally required for a group with that capacity. The moral failure is precisely the non-existence of the plural subject. It is analogous to Michael’s wrong in section 3.

There is an important difference between these two types of failure. The first involves the existence of a *plural moral subject* to which certain obligations are directly attributable: The group ought to do something; there is an obligation shared by the group. In the second, there are no obligations other than those of individuals. These interpretations are not contradictory alternatives. Rather, they are two ways in which a group of people could fail to do what they ought to do, morally speaking. The difference is whether “they” refer to the *collective* as a unit or as the *aggregated but separate* individuals.

## 5. Plural subjects

For the set-of-acts claim to be true, it must be possible for a set of acts to be wrong *as a set*. I have claimed that sets of acts by a person over time can be wrong as a set because the person has obligations over time, which include obligations to perform or refrain from sets of actions. I have suggested that we can understand the obligation of groups of people in the same way. An act can be right or wrong because it is one of a set of acts that a group ought to do or not to do. For this claim to be non-circular, the obligation of a group must be something more than just an aggregation of the individual obligations the members of the group have for independent reason. We are interested in cases where the obligations of individuals derive from the obligations of the group *as a group*, like the obligations of an individual at a given point in time derives from that person’s obligations over time. In this sense, what does it mean to say that a group ought or ought not to do something?

To answer what it means for a group that it ought *as a group* to do something, we need to clarify what it is for a group *to do something as a group*. In the individual case, we have a concept of personal agency that unifies the intentions and actions of successive self. What is the equivalent in the interpersonal case?

We often speak of *groups* of people intending things and doing things. For example, my family intends to go on vacation to Italy this year; and the United States invaded Iraq. We

find it natural to use a collective word as the subject of such sentences. This is also true when we talk of obligations and related moral concepts. We find it natural to say that we ought to do something. This is what the set-of-acts claim is fundamentally about. It is a claim that a set of acts, *by different people*, can be the right or thing to do. It is a claim that there are some things that *groups* should or should not do. If some version of the set-of-acts-claim is true, then there are cases of obligations of groups that are logically prior to obligations of individuals. In these cases, the rightness and wrongness of what individuals do derives from the rightness and wrongness of what groups do. To understand how this could be true, we need to understand what it is for a group to do something. We need to understand the concept of a *shared action*. And since the plausible accounts of group action are based on the concept of *shared intention*, I shall here roughly sketch an explication of that concept.

Plausible concepts of shared intention and shared actions avoid two mistakes. One mistake is to think that it is reducible to individual intentions about their *individual* actions, supplemented only with an adjustment of one's own actions in the light of how others are acting. This is not to deny that there are many situations where we naturally say that "we" or "they" are doing something without meaning anything more than that each of the individuals making up the group are doing that thing. When I tell you how people travel from New York to Philadelphia, I can tell you that "they usually take the train" without meaning anything more than that most people, most of the time they travel to Philadelphia, they intend to individually take the train, and they individually act on that intention. They adjust their actions to those of others—they decide when to go to the station depending on when the train driver and conductors will get the train there, for example—but their intentions do not *refer* to the actions of these others in any way different than would take into account physical constraints on what they could do. So their intentions are exclusively about their own actions—they are "I-intentions". That is not so in the much-discussed example of *travelling together* with someone. When two people travel together—travel *as a group*—that action cannot be reduced to individual "I-intention". When you and I travel together, it matters to me not just that we physically travel together, but in part that you intend to travel *with me*. My intentions in this case are "we-intentions"—they are intentions about what we do together. So the pronoun "we" play quite different functions in the two cases:

- "We usually travel by train" (said by a commuter from New York to Philadelphia, about the group of people who commute from New York to Philadelphia)
- "We are going by train" (said by one of a group of people who are taking a trip to Philadelphia together)

In the first case, the "we" is *aggregative*; in the second, it is *collective*. The same distinction applies to other terms for groups of people. (We may use "collective" intention or action synonymously with "shared" intention or action.)

So the first mistake to avoid is to think there is no sense in which acting as a group is different from acting as an aggregation of individuals. Shared intention and shared action cannot be understood in terms of aggregative terms only. But we must also avoid the opposite mistake of thinking that when a group acts as group, the group is something more than the individuals that make it up. While collective terms are necessary, they do

not refer to ontologically separate entities. A collective intention is not one that is held in a collective mind. There are no collective minds; only individual minds.

An adequate account of shared intention or action must therefore describe them as consisting of certain *individual* attributes, but these attributes must involve “we-intentions”. Different such accounts have been proposed. Michael Bratman defines a shared intention to *J* (in a two-person case) thus:<sup>7</sup>

We intend to *J* if and only if:

1. (a) I intend that we *J* and (b) you intend that we *J*.
2. I intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b; you intend that we *J* in accordance with and because of 1a, 1b, and meshing subplans of 1a and 1b
3. 1 and 2 are common knowledge between us.

And we can define *J* as our shared intentional action when we *J* as a result of 1, 2, and 3.<sup>8</sup>

This schema succeeds in explaining “what it is for us to intend something primarily in terms of (a) intentions and other attitudes of each and (b) the relations of these attitudes to each other.”<sup>9</sup> So does Margaret Gilbert’s account in *On Social Facts*. Her concept of “joint readiness” plays a similar function as Bratman’s joint intention:

[*A*] set of persons are jointly ready to share in action *A* in circumstances *C* if and only if it is common knowledge among them that they have mutually expressed their quasi-readiness so to share.<sup>10</sup>

where quasi-readiness

“involves a *conditional commitment of one’s will*, made with the understanding that if and only if it is common knowledge that the relevant others have expressed similar commitments... the wills are together committed or dedicated to the pursuit of the goal in question in the circumstances.”<sup>11</sup>

Like Bratman, Gilbert sees shared intention or commitment to act together as a matter of interlocking individual intentions or commitments.<sup>12</sup> I shall use Gilbert’s notion of a

---

<sup>7</sup> Bratman, “Shared Intention,” *Ethics* 104(October 1993), 106.

<sup>8</sup> Bratman defines a *shared cooperative activity* as such a shared action *J* which in addition to conditions 1–3 satisfies certain requirements of non-coercion and “cooperative stability,” essentially some willingness to help achieve the success of the *other* person’s intentions in 1–3.

<sup>9</sup> Bratman, 102–3.

<sup>10</sup> Gilbert, *On Social Facts*, 197.

<sup>11</sup> Gilbert, *On Social Facts*, 198.

<sup>12</sup> Anna Stilz argues that the difference between Bratman’s and Gilbert’s view is that the latter conceptualizes sets of people in a metaphysically suspect way, as ontologically independent entities: “Gilbert distinguishes herself from Bratman and Searle by arguing that in order to explain collective action, we must postulate an entity distinct from individuals which holds the intention to act. In consequence, on her view, that it is not sufficient for an individual alone to renounce a shared goal in order for the shared intention to cease to exist... This is because only the *plural subject as a whole* holds the intention... Since individual intentions are not the locus of explanation for collective action, for Gilbert, no change in individual intentions is sufficient to explain why collective action should or should not cease to occur” (Anna Stilz, *Solid Citizens*, unpublished book manuscript, 2007).

“plural subject” to indicate that these conditions on individual attitudes obtain. The term should be understood in an entirely non-metaphysical sense. That is, a plural subject just *is* the presence of the attitudinal states requisite for shared intention in two or more people. On this view, there is a plural subject whenever there is a shared intention to *J together*. We may say in these cases that it is the plural subject that intends to *J* (“we intend to *J together*”). This is analogous to saying that there is a unified subject John when John-1, John-2 have the appropriate attitudes towards each other, and that we can ascribe intentions to John “as a whole.”

I may have an intention to *J*. I may have the more general attitude of having a goal *X*. If it is my goal to achieve *X*, I am likely to have the intention to “what it takes” to achieve *X*. More strictly, I may have the intention to perform whichever actions may turn out to be necessary to achieve *X*, within certain constraints. Having the goal *X* does not mean I have an intention to *J* even if *J* is necessary to achieve *X*, if *J* is also going to frustrate some other goal *Y* that I also have and that I value more highly than *X*. But if I had no intention to do *J* if *J* was necessary to achieve *X* and in no way went against other values, I could not plausibly be said to have goal *X*. Since one could have as one’s goal simply to perform an action, my holding a specific intention to *J* is just a special case of my having a goal *X*—namely, the goal that I *J*.

Along the same schema as having shared intentions a set of people could have a shared goal. There is a plural subject whenever there is a shared intention together to perform whichever actions may be necessary to achieve a certain goal, within certain constraints. We may say in these cases that *X* is the goal or purpose of the plural subject (“we aim to

---

I believe Gilbert’s view is closer to Bratman’s than Stilz allows. Gilbert says: “Human beings X, Y, and Z constitute a collectivity (social group) if and only if each correctly thinks of himself and the others, taken together, as ‘us\*’ or ‘we\*’” (Gilbert, *On Social Facts*, 147). This is plainly a definition in terms of individual mental states. Her technical definition of the first person plural pronouns that people must correctly use, moreover, does not refer to any “group mind” but to a set of conditions on the intentions and attitudes of the individual minds of the members of the group, as well as conditions on the relationships between them. She says, for example, that the pronoun “refers to a set of people each of whom shares, with oneself, in some action, belief, or attitude, or other such attribute” (168). Thus her definition refers essentially to individual intentions. It is true that she thinks the necessary intentions must be we-intentions of a certain type and must themselves refer to concepts that are ultimately reducible to notions of individual agency only. This seems quite compatible with the analysis of shared intention and shared cooperative activity offered by Bratman, who allows that “[his] claim is not that we can specify [the contents of individuals’ attitudes] in ways that do not appeal to elements outside the individual whose attitudes are in question” (Bratman 1993, 113).

One exception both Bratman and Stilz takes to Gilbert’s analysis is her discussion of the rights and obligations that “come with” collective agency. They both argue that the fact of shared intention *by itself* does not necessarily generate a moral obligation to perform one’s part if one renounces the intention. But Gilbert explicitly says she is not discussing moral obligations: “The rights and responsibilities most directly in question [that one takes on or accepts when joining a group] are neither moral nor legal. They could be referred to as ‘associational’, or, more familiarly, as political... The deep basis for the rights and responsibilities that accrue to one who enters a group is to be found precisely in the concept of a plural subject, the core conception here. (As I have stressed, neither moral nor legal responsibility is at issue.) My articulation of the concept of a plural subject enables us to explain how membership in a group can entail a set of *sui generis* responsibilities and rights, and to become clearer about the nature” (Gilbert 411, 413; see also 162). I read Gilbert as discussing certain attitudinal states of commitment to act and expectation that others will act in certain ways, which attitudinal states are, according to her, part of what constitutes collective agency. Whether collective agency involves *moral* rights and obligations is a different question.

achieve *X* together”). In the case of shared goals, and in particular goals that require complex coordination of behaviour, plural subjects are often institutionalised. The set of people who incorporate a business corporation, for example, share an intention to act together in certain ways (pool a certain amount of resources and manage them together) so as to make profits. It is natural to say in these cases that *the corporation* intends to make profits and that it intends to act in certain ways to achieve that goal.

A set of acts is wrong if the subject to whom the set of acts “belongs” is acting wrongly by performing the action that the set of acts constitute. When is a moral subject acting wrongly? That question is the topic of the bulk of our moral thinking, as far as the individual case is concerned. Our interest here is in *plural* subjects, which unify, in the rough sense sketched above, multiple individuals through interlocking intentions. When, then, is a plural subject acting wrongly? Here are two ways of acting wrongly: By doing something one ought not to do; and by failing to do something one ought to do. One can do wrong, that is, by violating negative duties or positive duties. Can plural subjects have these duties?

If actions can be attributed to plural subjects as such, then there is no reason not to attribute negative duties to plural subjects as such. Negative duties take the form: “Subject *S* ought not to do action *X*” or more generally “No subject ought to do action *X*.” Collective actions could be wrong because of the way they treat or affect other people, just like an individual’s sequence of actions over time can be wrong for those reasons. Negative duties have to do with things it is not morally allowed to do to people—for example, killing them. It is what these actions do to people—such as killing them—that makes them wrong, regardless of *who* is doing the action. Whether a set of actions violates negative duties therefore does not depend on whether it was performed by an individual or a plural subject. Consider

*The Stoning.* A group of people resolves to stone a person to death. Their religion prohibits any individual from killing another. So they make sure to only use moderately sized rocks. There is no single killing blow; no one rock ends the victim’s life. Indeed no one rock makes a difference to the victim’s suffering. Still the victim dies a painful death.

*The Solo Stoning.* An individual resolves to stone a person to death. He makes sure to only use moderately sized rocks. There is no single killing blow; no one rock ends the victim’s life. Indeed no one rock makes a difference to the victim’s suffering. Still the victim dies a painful death.

In the Stoning, do any of the individual stone-thrower kill the victim? They do not. We should say that *the group* kills the victim, or that they all kill the victim *together*. We should also say that *the group* does something wrong, or that they all do something wrong *together*. The wrong that they do together is exactly the same wrong as the Solo Stone-thrower does. The morally wrong sets of acts are the same in both cases, and they are wrong because they violate the negative duty not to kill. Both the group and the Solo Stone-thrower have the ability to kill the victim. That is enough to give them both a duty not to. Thus the wrong committed by the group is immediate, it is not a function of any wrong committed by the individuals separately. This is not to say that the wrong does not “fall on” the individuals—it does, as the group has no independent existence from the

individuals who make it up. But the wrong committed *together* is logically prior to any wrong attributable to individuals.

We can ask one more thing of the claim that plural subjects have negative duties. I have said that the group, when understood collectively, has a duty not to kill. Ought implies can. So it must be possible for the group not to kill. One might say: “This is not always possible. A plural subject is set up by the interlocking intentions of its members. Suppose this group became a plural subject just because of individual’s interlocking intentions to kill the victim by stoning. Then the plural subject just *is* those interlocking intentions. *It* couldn’t act differently. Perhaps it could dissolve and reconstitute itself for other intended shared activities. But then it would be a different plural subject. So the plural subject in *The Stoning* cannot have this obligation, since ought implies can, and it cannot not kill the victim and still be this same plural subject.” This argument is mistaken. Usually a plural subject can intend to do something different; it can change its intentions, adopt new goals. But even if not, that does not imply it cannot avoid breaking its duty. A negative duty is avoidable by no action being taken—also when that is because there is no subject that can take action.

What is true is this. A group *Js* when its members do their parts of what constitutes the group’s *J*-ing, as a result of a shared intention to *J*. So a group can only violate its negative duties through intentional action. It is important to be clear what this claim is. A group can be acting wrongly even if it does not intend to violate negative duties. It will be acting wrongly if it intentionally acts in a way that, intentionally or not, ends up violating a negative duty. But it must intend the act in order for the *group* to be acting together. Without shared intention, a set of acts (which do not individually infringe on anyone’s negative rights) cannot violate a negative duty, because there is no subject whose duty it is. Without shared intentionality, the Stoning is like the Stock Market Crash.

Can plural subjects also have positive duties? Positive duties are of the form “Subject *S* ought to do action *X*.” Someone has a positive (moral) duty to perform an action if this someone can perform it and it would be morally wrong not to do so. So a group can only have a positive duty *as a group* to do something that *can be done together*. A duo could not have an obligation to sing a solo. (It could perhaps have an obligation to ensure that one of its members sing a solo, but that is not the same as an obligation to sing a solo.) More generally, it only makes sense to say that a group is doing something (rather than that an aggregate of individuals are each doing something) if it is appropriate to say that the group is a plural subject. The only actions a plural subject can perform are shared actions. Thus its positive duties can only be duties to perform certain shared actions.

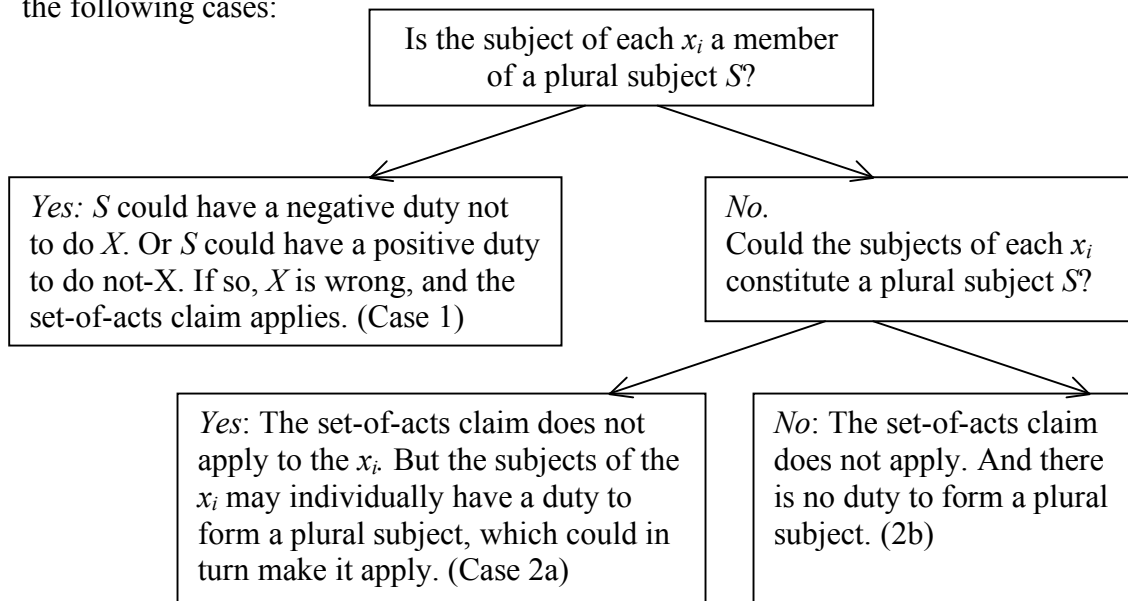
At the same time, an individual *qua* individual cannot perform a shared action. She can at most perform her part of a shared action, and when she successfully does so, she is part of a plural subject that performs the action in question. It is only *qua* member of the plural subject that she can be described as *e.g.* “playing a duet.” *Qua* individual she can only be described as “performing the actions that constitute her part of playing a duet” which may or may not actually involve her being a member of a duo that is playing a duet. (She may be hallucinating that there is another player.) This explains why the content of obligations of groups must be different from the content of the obligations of individuals.

To violate a positive duty to do  $X$  a subject must do not- $X$ . It is not sufficient for a positive duty to be violated that  $X$  *not be done*. For the reason why  $X$  is not done could be that there is no subject who could do  $X$ . In that case, there is no subject whose positive obligation it could be to do  $X$ , and so there is no obligation the non-doing of  $X$  could violate. If  $X$  is a collective action—a set of acts performed by a group *as a group*—then there may not exist a plural subject that could be the subject of the action. Given the intentionalist definition of plural subjects, when a group of people *do not do anything together*, they do not constitute a plural subject. (The exception to this is when a group is jointly ready to do something in certain circumstances but the circumstances do not prevail, or if it is not yet determined what has to be done to achieve its shared goal.) A non-existent plural subject cannot have duties. But shared actions can only be performed by plural subjects, so the actions that groups *could* have a positive duty to do, cannot be actions individuals not constituting a plural subject could have a positive duty to do. A group of people who *could, if they constituted a plural subject, together achieve morally valuable things and have a positive duty to do so*, do not by that fact constitute a plural subject whose non-action violates those duties. Without plural subjecthood, the set of acts which constitutes the lack of action *as a group* is not, as a set, wrong. Hence the set-of-acts claim does not apply. This is true in the Stock Market Crash.

One important implication of these last paragraphs is this: A set of acts by different people cannot be wrong if there is no plural subject whose action they constitute (unless the acts are individually wrong). So the set-of-acts claim does not apply when there is no plural subject constituted. We may refine the generalised set-of-acts claim thus:

Even if an act is not wrong by itself, this act may be wrong because it is one of a set of acts that together constitute a wrongful action (or inaction) by an existing plural subject. Similarly, even if some act is not what someone ought to do in isolation, it can be what someone ought to do, because it is one of a set of acts that together constitute the action that ought to be done by an existing plural subject.

We can now summarise what we have claimed so far. Consider a number of individual actions  $x_1, x_2, et\ cetera$ . They may be performed by the same individual at different times, or by different individuals. Denote by  $X$  the set of all the actions  $x_i$ . The set-of-acts claim is that one or more of the  $x_i$  could be wrong because  $X$  is wrong. We have distinguished the following cases:



In case 1,  $X$  is a wrongful action by the group  $S$ . This is analogous to John's not watering the plant. In this case, the set-of-acts claim applies. The individual acts  $x_i$  *could* therefore be wrong. I discuss in the next section when they are.

The second case is when the people whose actions make up the set of acts  $X$  do not form a plural subject, and so they cannot be jointly doing  $X$ . We further divide this case according to whether they (2a) *could* be a plural subject or (2b) not. If they could, then the second case (2a) is analogous to Michael's not doing  $X$ . If they could not, then the second case (2b) is analogous to Leonard's not doing  $X$ . I discuss these cases in section 7.

## 6. Existing plural subjects

### *a. Members' responsibility for collective obligations*

I have claimed that the set-of-acts claim applies when it is meaningful to talk of a subject that is responsible *for the set of acts as a whole*. We can draw on accounts of shared intentions, shared actions, and plural subjects, to delineate when that condition holds. When it does, the subject can have a negative duty not to perform certain sets of acts, and a positive duty to perform certain others. These possibilities mean that a set of acts can together be wrong, so the set-of-acts claim applies. We must now ask how it applies. *When* is an act wrong because it is one of a set of acts which together are wrong?

Consider first negative duties. When a group acts as a group, the group can harm people or violate their rights. When it does, it acts counter to a negative duty. In the Stoning, the group fails to fulfil its negative duty of not killing others. We have assumed that the group doing the stoning is a plural subject—all its members intend to do the stoning *together*. Their goal is to *collectively* kill the victim, while not individually killing the victim. They achieve this goal. They together violate a negative duty. They do not do so *qua* individuals. No individual kills the victim. Yet every individual, *qua* member of the group, does. *As members of the group* each individual violates the negative duty.

Could a member claim that *he* is not acting wrongly, since *his* rock does not make a difference? He could not. He *is* acting wrongly. Indeed he would be acting wrongly even if he did not throw any rocks, as long as he is a member of the group. This is clear from

*The Disabled Stoning Participant.* One of the most vocal members of the group lost both arms in an accident several years ago. He will not be able to throw any stones. Yet he takes part in the group's deliberation and decision to carry out the stoning. He comes along to the stoning and cheers, although he throws no stones.

The disabled stoning participant violates a duty not to kill just as much as the able stone-throwers. *Qua* member of the group, he kills the victim (or takes part in killing the victim). He does not inflict any harm on the victim individually. Nor do the able stone-throwers. They do not kill the victim except as members of the group, and in this they are morally equivalent to the disabled stone-throwing participant. The Disabled Stoning Participant's situation is similar to that of John once the plant he failed to water is dead. He is not on that day acting wrongly; nor is he individually acting in the same ways as his

earlier selves (they failed to water a *live* plant); yet he is at all times responsible for the failure over time to care for the plant.

Violating negative duties is a wrong of commission. If a group can act as a plural subject, then it can commit an act that is wrongful by commission. In this case, the members of the group that is acting wrongly are acting wrongly *qua* members of the group. This could be the case even if they are not doing anything that individually involves wrongdoing, as in the case of the Disabled Stoning Participant. We may say more generally:

*Principle of Collective Responsibility (PCR)*. If a group of people are acting wrongly *together* (i.e. as a plural subject), then each member of the group is acting wrongly *qua* member of the group, regardless of whether he is acting wrongly as an individual.

PCR is at the same time a weaker and a much stronger claim than Parfit's principles.<sup>13</sup> It is weaker because it requires that the group act *together* in the shared-action sense already outlined. It is stronger in two ways. It is stronger because it does not require that the outcome be *worst* if *all* throw rocks (because a wrongful action by the group need not require every member to do anything individually), only that the outcome be *worse* if *enough* throw rocks. More importantly, it is stronger because it says each member is acting wrongly, even if he is not himself performing the actions that directly constitute the killing. (*A fortiori*, it says that the Skeptical Torturer is acting wrongly in torturing together with the other torturers.) This may seem to violate the precept of "ought implies can." How can *I* be acting wrongly because the group of which I am a member is acting wrongly, if there is nothing I can do about what the group is doing? But the point is that the *group* is acting wrongly, and so it is at the level of the group that "ought implies can" is required. Since the shared action is shared by *each*, so is the wrong.

Recall the case of Watering Plants. When the end of the summer approaches, John may protest against an accusation of wrongdoing. He may point out that there is nothing he can do on the last day of the summer to save the plant, given that his earlier selves have failed to water it. How can the later John be acting wrongly because John-over-time acted wrongly, when there is nothing the later John can do about what John-over-time has done? But the point is that John-over-time has acted wrongly. And we can attribute this wrong to any of the successive Johns, for example by requiring a later John to replace the dead plant. Similarly we can attribute the killing by the group in the Stoning to every member of the group.

This argument may seem to have the undesirable implication that we cannot say of any of the stone-throwers that they have an obligation not to throw any rocks. If it did entail that conclusion, we would have to reject it, for that conclusion is absurd. But it does not entail that conclusion, for two reasons. The first is that it ignores a special case. If only *one* member of the group carries out the stone-throwing (or more generally, if few enough do that each one makes a difference), then he *can* control what the group does, and he ought clearly to act in such a way that the group not violate its negative duty. So he ought not to throw his rocks. But a special case does not answer the general charge. The second reason why the argument does not entail an absurd conclusion is this. It only entails that

---

<sup>13</sup> The relevant principle is the one stated in footnote 5.

if a member ought not to throw rocks, this cannot be because not throwing rocks by itself means he would not be acting wrongly. This is what the Disabled Stoning Participant shows. But there could be *other* reasons why a member ought not to throw rocks.

The reason could be this. Only by sharing in the collective action is a member acting wrongly. The member shares in the collective action by being a member. So if he could *leave* the group before it has violated its negative duties, he would not be acting wrongly, since he could only have been acting wrongly *qua* member of the group. So if possible, he ought to leave the group. It is a matter of some dispute exactly what is required to leave a plural subject of which one is a member. We shall not here address that question. But it seems clear enough that it must *at least* involve refraining from participating in the collective action through which the group violates its duty together. So if it is possible to leave the group, each member ought to do so, and as part of leaving (or as a consequence of it), ought not to throw rocks. The stone-throwing member therefore acts wrongly in two ways

Now consider positive duties.

*The Unreliable Fire Fighters.* I am a member of a fire brigade. The fire brigade has a positive duty to put out fires (perhaps they have promised the community to do this; perhaps they have a unique ability to do so). My job is to man the fire engine. Others are charged with operating the hose. One day there is a fire in the neighbourhood, and I am at the engine. The other members fail to turn up.

The fire brigade *together* hold an obligation to put out fires. I share in this obligation: I have a positive duty to put out fires together with the rest of the fire brigade. So do the other members. Have I fulfilled this duty? I have not. Even though I have done my part—I am at the engine when the fire breaks out—we have not fought the fire. So I have not fought the fire together with the others, as was our obligation. I am like Michael on a day when he happens to water the plant, but who does not do so often enough to save it. Just like doing what is required on *one* day only does not discharge Michael's obligation, my doing *my* part only does not discharge *our* obligation to fight the fire.

So like in the case of negative duties, *qua* member of the group, I am acting wrongly regardless of whether I am doing my part. PCR applies. We ought again to ask whether this entails that we cannot say I have an obligation to do my part of the collective action that would discharge the positive obligation. It does not. There is a special case in which this does not follow. If my action is *pivotal*, so that we will fight the fire if I man the engine but not if I do not, then my action determines whether we are acting wrongly as a group, and consequently whether I am acting wrongly *qua* member of the group. In this special case, I ought to man the engine.

In the general case where enough others do not turn up that we will fail as a group to fight the fire whatever I do, however, manning the engine does not make a difference to fulfilling our obligation to fight the fire together. So *that* could not be the reason why I ought to man the engine. In the case of negative duties there was another reason. It was that I could avoid wrong-doing by leaving the plural subject, and this would involve at a minimum to refrain from my part of the set of acts that together violated the negative duty. Does this reason apply in the case of positive duties? It may not. For negative duties, a member does *not* share in the performance of actions that violate them until

those actions are committed. So they could leave the subject without having shared in the wrongful action. In the case of positive duties, in contrast, a member *does* share in their *non*-performance until the actions fulfilling them are committed. It takes more to “get out of” moral responsibility for a positive obligation that is currently unfulfilled than moral responsibility for a negative obligation that is currently fulfilled because it has not yet been violated. (Moreover, it would be odd to construe manning the engine when nobody else turns up as a necessary part of *leaving* the fire brigade.) Thus the obligation of the group does not, in the case of positive duties, seem to translate into an obligation for me to man the engine if I know that other members are not turning up.<sup>14</sup>

This may be thought to prove too much. Because it also seems that it does not translate into an obligation to man the engine if the other members *are* turning up, unless my action is pivotal. Consider:

*The Oversubscribed Fire Brigade.* Due to the popularity of fire fighters, our fire brigade has had many new and eager recruits. There are always enough fire fighters present for the brigade together to fight any fire that may break out. If I did not turn up, the brigade would, as a group, fight the fire just as well as me, although it would take each member more effort, as one of them would have to drive the engine and one fewer would be manning the hoses. I do not turn up for the next fire.

It seems intuitive that I have an obligation to turn up. My argument says that if I do, it is not because of the obligation the fire brigade has to fight fires. This is the right conclusion. My contribution does not affect the degree to which the group fulfils its duty. If I am acting wrongly in not turning up, it cannot be because of the group’s obligation.

### *b. Members’ duties towards one another*

Many think there is an obligation to “do one’s part” when enough others are doing so. Parfit believes this. McMahon believes rationality requires it. Their principles give

---

<sup>14</sup> An objection to this view is similar to the one discussed in the context of negative duties. It is that the argument implies that those who do and those who do not do their part of what would constitute the group’s fulfilling its obligation are acting equally wrongly. This would be an absurd conclusion. But it does not follow. Consider:

*The Taxi Ride.* Two strangers are together waiting for a bus to the airport. The bus never arrives, and they agree to share a cab. When they arrive at the airport, person 1 realises he has no cash. Person 2 has enough cash for three-quarters of the fare, and pays the driver all he have. As they leave the cursing driver, they realise they must run to catch their respective flights to different parts of the world. They do not have time to get each others’ contact information and never see each other again.

The pair have together failed to fulfil their duty to pay for the full ride. Have they acted equally wrongly? My argument does not entail this conclusion. It entails is *one way* in which they are acting equally wrongly. This is true. They have all equally failed their *joint* obligation *to the driver* to pay him in full. But Person 1 has additionally failed to do his part. As I argue below, this is failing an *individual* obligation to Person 2—an obligation to *do one’s part*. Therefore Person 1 is acting wrongly in more than one way. To see this, suppose Person 2 had had enough money to pay the whole fare. Then the obligation to the driver would not have been violated. Person 2 would not in any way have acted wrongly. Person 1 would still have acted wrongly—towards Person 2—in not doing his share. The wrong of not paying the *driver*, in the original case, is an *additional* wrong, and that wrong both people shared equally.

individuals reasons to participate in collaborative actions when enough are doing so to make a difference, even if the individual's own contribution makes no difference. For example:

*The Bucket Brigade.* The fire brigade has no fire engine, and relies on a bucket brigade to bring water to burning houses. During most fires, there are more fire fighters present than are needed to cover the distance from the nearest well to the burning house. During every fire, I always look around to check there are more than enough other fire fighters present, and if there are, I make sure to linger behind until the line is fully formed. (I do not have the training to do any other firefighting tasks, and there are enough trained fire fighters to fully carry out those tasks.)

Am I acting wrongly? Many people feel that I am. But we have shown that this cannot be because of the obligation the group has, since that is in any case fulfilled if enough other participate. The common sense moral intuition for collaborating in such cases is that doing otherwise is to “free-ride”—it makes others bear the costs of collaboration in fulfilling a shared obligation. That is, it is an intuition that individual members of the group owe something to *each other*. This suggests another principle:

*Principle of Solidarity (PS).* Members of a plural subject owe one another a duty to do their part of the set of acts that constitutes a collective action which they share either an intention or an obligation to perform together.

Why and when should we think the PS applies? The question of whether shared intention *ipso facto* involves certain moral obligations has been a matter of some dispute.<sup>15</sup> If it does, that would also cover a shared intention to fulfil an obligation. Here I will only briefly sketch some possible arguments to that effect. First, the common knowledge of shared intentions will normally create expectations that other members intend to do their part. It is conceivable that a joint intention could be formed and joint action performed where all the members publicly “reserve the right” to withdraw whenever they may wish (Bratman's case of “unhelpful singers” is a good example; see footnote 15). Without this, however, common knowledge of shared intentions usually involves a mutual encouragement of certain expectations. As Scanlon explains, leading others to expect a certain behaviour from one can generate a moral obligation to perform that behaviour.<sup>16</sup> Partly this is because the expectations created make others adjust *their* behaviour in light of the intended joint activity. Indeed the more extensive these adjustment are, the more

---

<sup>15</sup> See footnote 12. Bratman argues against this view. He says, for example:

Imagine two singers who each highly value their duet-singing but nevertheless have a clear understanding between them that neither is making any binding promise to or agreement with the other concerning their singing. Each publicly states that she reserves the right to change her mind. These two could still share an intention to sing a duet together. They could still engage in coordinated planning aimed at their singing the duet and in which each relies on the participation of the other. Granted, the normal case of shared intention will not be like this. In a normal case there will likely be some promise or agreement; and that will further contribute to the confidence of each that she can plan on the participation of the other. Nevertheless, such a promise or agreement does not seem essential to shared intention. And when there is no such promise or agreement, or some other obligation-generating process, the shared intention may not impose a nonconditional obligation to stick with the joint action.” (Bratman 1993, 111)

<sup>16</sup> Thomas Scanlon, “Promises and Practices,” *Philosophy and Public Affairs* 1999.

likely the joint activity affords its members something they could not achieve on their own. This in turn gives the members reason to generate the appropriate expectations in others, which increases the promise-like obligation to fulfil those expectations.

There is a different way intentions could in themselves involve moral obligations to do one's part. I mentioned earlier the distinction between single intentions and broader goals. A plural subject could be created to perform a single shared action, such as going for a walk or dancing together. More relevant for this discussion are plural subjects created to perform broader goals. These are formed by shared intentions to perform whichever actions are necessary to achieve the goal, with these left unspecified at the outset (within certain limits). A shared intention to do whichever unspecified shared action is necessary to achieve the goal must be construed as a form of *delegation*. It commits the members together to the actions that may be found to best achieve the goal at some future time, to be decided by some decision-making procedure. This shared intention, that is, must include the intention to decide the details according to some procedure. The delegation of decision-making power already constitutes a degree of consent to performing one's part of what the mechanism will later adopt as the appropriate collective action. It is also more likely to create expectations that will lead to large shifts in behaviour. For the same Scanlonian reasons, this strengthens the moral reasons for doing one's part.

A very different justification for the PS comes from our discussion of the individual share of collective wrongdoing earlier in the section. I showed there that when the a plural subject acts wrongly, all the members equally are acting wrongly *qua* members, regardless of whether they are doing their part of what would be the right collective action or not. The fact that the group has a collective obligation means that some division of the burden of complying with has to be allocated somehow—in other words, some of the members have to do the things that make the group fulfil its obligation. Now typically those who do not do their part gain some private benefit (avoid a private cost) from doing so. We may argue, therefore, that those who do not do their part impose an unfair balance of burdens and benefits on the rest of the group.

Note that the claim is not that they *increase* the burdens of the members of the group overall, since that is only the case for individuals whose actions are pivotal. It is, rather, that they are unfairly availing themselves of greater benefits without taking a greater share of the burden. Thus an internal moral requirement of justice between people who share collective obligations could be seen as justifying the PS.

These brief accounts have two desirable conclusion in common. First, they both give a reason why members of a plural subject which is together doing what it ought to do, should do their part. The reason is an obligation to their fellow members. They have either committed to doing so in a promise-like fashion, or they would be imposing an unfair burden on others if they do not. Second, they show why it is plausible to say that when others do not do their part, so the plural subject is not overall doing what it ought to do, then there is less reason to do one's part. For the only reason why one should do one's part is the duty one has to fellow members to do so—but if the plural subject is not performing the action it ought to do, that means not enough other members are fulfilling their duty to fellow members. Surely we have less strong reason to fulfil duties towards those who violate those very same duties towards us, than towards those fellow members

who faithfully do their part. Thus in *The Unreliable Fire Fighters*, my only obligation to man the engine, when nobody else turns up, is my obligation of solidarity to those other fire fighters to do my part. But when they have violated their exact same obligation towards me, the force of the obligation for me is slight, and can justifiably be outweighed by other considerations.

## 7. Putative plural subjects

On the account I have offered so far, there are no individual obligations that derive from the mere fact that a set of acts would together have morally valuable effects, unless the set of acts can be attributed to a unified subject. That is not to say there is nothing regrettable about situations where individuals fail to act collectively when doing so could produce morally valuable outcomes. Such situations are very regrettable. But on my account, they do not involve wrongdoing. They are regrettable in the sense accidents or natural disasters are regrettable. They cause morally significant losses, but do not involve wrong-doing.

Is this a problem with the account? Some think that even in the absence of plural subjects, individuals can be acting wrongly if they don't *instigate* collective action. Suppose that there is no fire brigade in my community. Am I acting wrongly if I do not attempt to form one? Parfit suggests that I may be:

“Suppose that someone has done the act, of those that are possible for him, whose consequence is best. It does not follow that this person has done what he ought to have done. He ought to have asked whether he is a member of some group who could have acted in a way whose consequence would have been even better. If this is true, and he could have persuaded this group to act in this way, this is what he ought to have done.”

So if I could make others act together with me—that is, if I could bring a plural subject into being—then, Parfit suggests, I ought to do so. Should we believe this? Whether we do or not, we should believe that if I *cannot* bring a plural subject into being, I am not acting wrongly. This is like Leonard who has anteretrograde amnesia.

If a plural subject is possible, we should distinguish between two possible cases. We could ask whether I ought to do what I can to *bring one into being*; and whether I ought to *join* a plural subject that already exists. Note that in general, bringing a plural subject into being is itself a set of acts by different people. So in general, bringing a plural subject into existence—“organising ourselves,” we might say—can only be “our” obligation, but in the absence of a plural subject, there is no “we” that is the subject of obligations. So any individual obligations to join or form a plural subject cannot derive from such a collective obligation.

What then could be the source of an individual obligation to join or form a plural subject? There is a special case. It could be that enough people are disposed to join in a shared action or intention *conditional* on everyone else joining—that is, conditional on a plural subject existing. I may be in a position to trigger those conditions, perhaps by making them common knowledge, or by “giving the sign.” Some individuals are pivotal in virtue of their particular position among others. Since their individual act can make the

difference between the presence or absence of collective action, they may have an obligation to “launch” a plural subjects for ordinary (individual) moral reasons.

In the general case, however, individuals cannot have an individual obligation to set up a plural subject, since this not something they can individually do. And I have shown that they cannot derive such an individual obligation from a collective obligation. What they could have an individual obligation to do, is to adopt the attitudinal disposition corresponding to collective action, conditional on others’ doing the same. This amounts to an obligation to *join* some plural subject if one exists. Why would individuals have such obligations? It could not be because it leads to better outcomes, because their individual participation typically makes little difference. It could not be because they have that obligation *qua* members of the group, because the issue at stake is whether they should become members of the group. Nor could it be an obligation of solidarity with other members of the group.

This situation is equivalent to that of Michael, the man who lives in the present. The individuals of whom we are asking whether they ought to join a plural subject are in the same position vis-à-vis the plural subject as Michael on any one day is vis-à-vis Michael-over-time. What *moral* reasons could we give Michael today to see himself as an instantiation of the continuing agency of Michael-over-time? I said earlier that we attribute to him a moral power to unify his successive selves into a continuous agent. But this means we attribute some moral power to him *as* a continuous agent. In light of the argument I have offered here, we should have doubts about this. Are we not simply positing a “superagent” who can bind the individual time-slice decision-makers? We may accept this. It may seem to us such a central part of our moral life that we must take the unity of agency within individual lives as a primitive. But then we have less reason to resist “collective superagents” who can bind individual members. For it could seem to us an equally central part of our moral life that we are members of groups. If we can require Michael-today to defer to the responsibilities of a posited Michael-over-time *that Michael-today himself rejects*, why can we not require an individual to defer to the responsibilities of a group that the individual herself rejects? We would of course have to give relevant criteria for establishing the unity of different selves, and they should be similar in the intrapersonal and the interpersonal cases. They might have to do with similarity, “being in the same boat,” ability of one self to affect the situation of others, and so on. This could prove difficult in the interpersonal case. But surely it could sometimes be done. In such cases, we might be led to positions with which we are uncomfortable. And we would still have to answer the metaphysical question of how we can posit unifying agents.

We ought therefore to revisit our original intuitions about Michael. The question is whether we can give him a moral reason to see himself as single agent unifying his successive selves over time. It may seem to fanciful to think that someone could not see himself as a single agent from day-to-day. But over longer periods of time, people do repudiate their earlier selves and the later selves they may become. So even if we find preposterous the idea that Michael repudiates his earlier selves every day, can we give him a moral reason to see himself as a single agent over the course of his whole life?

It seems to me we cannot give him (them?) a moral reason. We *can* give him prudential reasons. We can show how each of his successive selves can be made better off thanks to

the planning and coordination unified agency provides. But he may not (today) care about this. Indeed if Michael-today is the first of the successive selves to “unify,” he is likely to bear much of the costs and reap few of the benefits. As Parfit has shown, it is not *irrational* to care only about one’s present aims. We may appeal to the inherent value of unified agency. Most of us do find value in the continuity of the agency that links our successive selves. Most of us find a similar value in acting together with others. Michael may be mistaken in not seeing this. But he may not be.

The nearest we can get a moral reason, I think, is to say that Michael’s refusing to “fuse” each successive selves into a continuous agent, reflects badly on his (or rather, their) character. He lacks integrity. He does not reach the depth of identification with a life project of which human beings are capable. This does not amount to showing that he is, at any time, acting wrongly. It does not generate an *obligation* for him to act differently. Perhaps this is a failure for morality. But as Bernard Williams has pointed out, there is more we can say in ethics than identifying obligations.<sup>17</sup> It may just be that the correct place for statements about what is ethically unsatisfactory in Michael’s behaviour *is* in the judgment of his character.

We might be able to say something similar about the group of people who never “get their acts together” to act as a plural subject. We may say that this manifests bad character in the individuals, and a bad culture within the whole group. But on this account, at least, we cannot say of any individual that he has a moral obligation to join a plural subject, and that he is acting wrongly in not doing so. If this seems to treat too lightly the ethical challenge of such situations, that may simply mean that we have to go beyond the language of moral obligation to describe the challenge.

---

<sup>17</sup> See especially chapter 10 of *Ethics and the Limits of Philosophy*.