

# **Our Obligations to Act Together**

Luke Devereux

University College London

MPhil Stud Philosophical Studies

## **Declaration**

I, Luke Devereux, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

## Abstract

In 1970 Virginia Held wrote a paper in which she asked the reader to imagine a fictional scenario in which some strangers become bystanders to a murder whilst riding public transport. Held was interested in questions of moral responsibility – specifically, could the group of bystanders be morally responsible for its failure to act?

This thesis uses Held's example as a lens through which to ask a series of questions about moral responsibility and blame, in the hope that the answers to these questions can provide insight into our real-world duties to act with others. Given that many problems facing society can plausibly only be overcome by groups of people acting together, it is important that we have a good understanding of how we are morally required to behave collectively.

After a short introduction to Held's example and some clarificatory remarks, the first chapter of the thesis is dedicated to an enquiry into the nature of moral responsibility and blame. The second chapter examines Held's proposal that we should hold the group *qua* group morally responsible for its failure to act. I argue that, whilst in other contexts it is appropriate to hold groups *qua* groups morally responsible for their acts, the bystanders in Held's example do not constitute a kind of group which is capable of being morally responsible. We can therefore only ascribe moral responsibility to individuals in Held's example. The third chapter asks how we can hold people morally responsible for failure to act when their action couldn't have made a difference. The final chapter focuses on how my findings might apply to our real-world obligations.

## **Impact Statement**

I anticipate that this thesis will contribute to the ongoing philosophical debates on collective responsibility, blame, group action, and the moral significance of symbolic gestures.

## Table of Contents

|   |     |
|---|-----|
| <b>Chapter 1: Held's Subway</b> .....   | 6   |
| <b>Chapter 2: Some Moral Concepts: Responsibility, Blame and Obligation</b> ..... | 17  |
| 2.1 Responsibility .....  | 17  |
| 2.1.1 'Responsibility' and its cognates .....                                     | 17  |
| 2.1.2 Hart's classification.....  | 18  |
| 2.1.3 Moral responsibility .....  | 22  |
| 2.2 Blame .....   | 31  |
| 2.2.1 Why ask the Blame Question?.....  | 31  |
| 2.2.2 Scanlon's account of blame.....   | 33  |
| 2.3 Obligation.....   | 36  |
| 2.3.1 When does 'can' imply 'ought'? .....  | 36  |
| 2.4 Conclusion.....   | 49  |
| <b>Chapter 3: Held's Answer</b> .....   | 51  |
| 3.1 Random Collections.....   | 51  |
| 3.2 Some Important Distinctions .....   | 53  |
| 3.3 Which Collective Responsibility? .....  | 60  |
| 3.4 Against Held's Answer .....   | 63  |
| 3.5 Conclusion.....   | 72  |
| <b>Chapter 4: What If I Couldn't Have Made a Difference?</b> .....                | 73  |
| 4.1 Case 1: When Others Might Also Cajole.....                                    | 74  |
| 4.2 Case 2: Five Bad People .....   | 85  |
| 4.3 Case 3: One Good Person, Four Bad People.....                                 | 86  |
| 4.3.1 Why the One Good Person might be blameworthy.....                           | 86  |
| 4.3.2 Some ethical theories which can't explain the intuition.....                | 89  |
| 4.3.3 An ethical theory which might explain the intuition .....                   | 94  |
| 4.4 Conclusion.....   | 102 |
| <b>Chapter 5: Why This Matters</b> .....  | 105 |
| 5.1 The Busy Road .....   | 105 |
| <b>Bibliography</b> .....   | 111 |

## **Chapter 1: Held's Subway**

In 1970 Virginia Held wrote a paper in which she asked her readers to imagine the following scenario.

*Assume that there are seven apparently normal persons in a subway car; none is acquainted with any other; none are sitting together. The second smallest person of the seven rises, pushes the smallest to the floor, and, in full view and hearing of the remaining five, proceeds to beat and to strangle his victim. If the remaining five persons do nothing for, say, ten minutes, at the end of which time interval the smallest person is dead, would we be able to make a valid judgment that "they ought to have subdued the strangler"? It is possible that no one of the five, acting alone, could have subdued him; it is extremely probable that action by two or more of the group to subdue him would have succeeded, with no serious injury to themselves; the group was not so numerous that, if each member had acted, confusion would have resulted.<sup>1</sup>*

This hypothetical, though realistic, scenario has been central to my thinking throughout the production of this thesis. The majority of what is to come has been written with this example in mind, and it has taken on such significance that it is worthy of a name of its own: I will refer to this example as 'Held's Subway'.

---

<sup>1</sup> Held (1970), p.476-477

The reason that Held's Subway has so persistently stayed with me since I first came across it is, I believe, partly due to its vivid and shocking nature. More importantly, though, is that it raises important ethical questions which deserve careful attention.

Two of the most pressing questions surrounding Held's Subway are, I think, 'who is responsible?' and 'who is to blame?'. I will refer to these questions as the 'Responsibility Question' and the 'Blame Question' respectively. Although commonly used interchangeably in everyday conversation, these questions are importantly different from one another, and must be distinguished. I will consider the question of exactly how they are different from one another in due course.

As soon as our attention is drawn to the Responsibility and Blame Questions, answers begin to suggest themselves. For many people, the first answer that springs to mind is that the *attacker* is responsible and blameworthy. After all, he is the one who actually performed the horrible act of beating and strangling the victim. Had he not done so, the victim would have had an uneventful subway ride and thought nothing more of it. So it is surely right to say that the attacker is both responsible and to blame – we don't need a philosophy thesis to tell us that. However, most people will also have the intuition that that is not the end of the story.<sup>2</sup> Indeed, Held sets up the case in such a way that we are invited to go further and ask whether some or all of the remaining five bystanders are also responsible, and are also to blame. Shouldn't the bystanders have done something? Somebody who is satisfied to lay

---

<sup>2</sup> By 'intuition', I simply mean, as Ezra Rubenstein puts it, 'a belief which is pre-theoretically compelling' (Rubenstein (2015), p.6).

responsibility and blame at the feet of the attacker and leave things at that is surely missing something of moral importance.

What should the bystanders have done? Here, again, an answer readily suggests itself: the bystanders are blameworthy for not wading in to the fracas and stopping the attacker from carrying out his assault. Each bystander could have done what I will call 'unilateral intervention': they could have charged in by themselves and attempted to apprehend the attacker. That none of them did so means that they each bear responsibility for the death, and are blameworthy.

However, Held rules such an answer out. 'It is possible', she writes, 'that no one of the five, acting alone, could have subdued him'. Let us strengthen this by stipulating that it is *certain* that no one of them acting alone could have prevented the attack. If any of them had unilaterally intervened, let us imagine, all that would have happened is that they would have been beaten up too, perhaps to death, and the original victim would still have died. Let us also assume that each of the bystanders knew as much. In this case, it is not right to say that the bystanders are responsible and blameworthy for not unilaterally intervening. It would be excessively harsh to require of somebody that, in order to avoid responsibility and blame for an assault, they have to unilaterally intervene, risking their own life, even when they know for certain that it is not going to make any difference.

Unilateral intervention may be required to avoid what Karl Jaspers called 'metaphysical guilt'.

In describing metaphysical guilt Jaspers writes the following.

*There exists a solidarity among men as human beings that makes each co-responsible for every wrong and injustice in the world, especially for crimes committed in his presence or with his knowledge. If I fail to do what I can to prevent them, I too am guilty. If I was present at the murder of others without risking my life to prevent it, I feel guilty in a way not adequately conceivable either legally, politically or morally. That I live after such a thing has happened weighs upon me as indelible guilt.<sup>3</sup>*

This is the kind of guilt that Jaspers lays at the door of the German people who lived at the time of the Holocaust.<sup>4</sup> It is sometimes called ‘survivor’s guilt’, and refers to the sense of responsibility one feels simply for continuing to live after another has died in your presence. To avoid metaphysical guilt, Jaspers says, one must do everything one possibly can to save the victim – in Held’s Subway, this means unilateral intervention.

However, metaphysical guilt is not what is at issue here. I am not asking what the bystanders should have done to avoid metaphysical guilt; I am asking what they should have done to avoid responsibility and blame. Though I have not yet attempted to give a detailed account of these two concepts, pre-theoretical understanding of them suggests that they present a lower bar, as it were, than Jaspers’ metaphysical guilt.

None of the bystanders, I maintain, is responsible or blameworthy for failing to unilaterally intervene. Rejection of this answer to the Responsibility and Blame Questions, though, leads

---

<sup>3</sup> Jaspers (2000), p.26

<sup>4</sup> Ibid. p.65-67

directly to another straightforward answer. What seems so objectionable about the bystanders in Held's Subway is that, as the attack is going on for a full ten minutes, they just sit there passively without so much as attempting to do anything to help. We are compelled to say that what they should have done *something* – but what? Assume, for the sake of the example, that it was known to all of the bystanders that attempting to remonstrate directly with the attacker was pointless. An option that *was* available to them, however, was to attempt to persuade *each other* to come together as a team, and collaboratively attempt to overthrow the attacker. Held tells us that it is likely that cooperative action against the attacker would have been successful. 'It is extremely probable' she says, 'that action by two or more of the group to subdue him would have succeeded, with no serious injury to themselves'. Again, let's shore this up by imagining that two or more of the bystanders working together would *certainly* have stopped the attack, with no injury to themselves.

I will refer to the action of any one of the bystanders standing up and attempting to persuade the others to join forces as 'cajoling'. The action of cajoling and being the only one of the group to do so I will call 'unilateral cajoling'.

Now let us assume that unilateral cajoling on the part of any of the bystanders would have been successful. That is, had any of the bystanders unilaterally cajoled, at least one of the other bystanders would have acceded to the request, and the attack would have been stopped. Assume also that any of the bystanders could have unilaterally cajoled without risk of harm to themselves – perhaps the attacker was too absorbed in his brutal assault to notice what the other bystanders were up to, so there was no chance of him turning on any of them

if they had started to cajole.<sup>5</sup> In this case, the answer to the Responsibility and Blame Questions is clear: each bystander is responsible and to blame for not cajoling. Had any of them done so, a competent group would have been formed, and the attack would have been prevented.<sup>6</sup> The moral requirement to cajole comes, as I will argue in chapter two, from a requirement, when one is part of a group capable of bringing about a good outcome, to do one's best to ensure that the group coordinates in the way required to bring about that good outcome.

If Held's Subway is as I have described it so far, there is no difficult philosophical problem. The answers to the Responsibility and Blame Questions are straightforward: the attacker is to blame, for obvious reasons, and each bystander is to blame for not cajoling, since had any of them done so, the attack would have been prevented. There is no collective action problem, or problem of group harm, since each person could have brought about a better outcome if they had acted differently from how they in fact did act.

To motivate an interesting philosophical problem, then, I will make a further stipulation, which Held didn't make, or at least didn't make explicitly. I will stipulate that *if one of the bystanders had attempted by himself to cajole the others into helping him, he wouldn't have been successful*; unilateral cajoling would not have worked.

---

<sup>5</sup> This particular assumption, along with some of the others I have made, may be unrealistic. However, I ask the reader to accept these conditions for now, as they help us to get to the interesting features of the case.

<sup>6</sup> Perhaps there needs to be an epistemic condition here as well. I will add that so long as each bystander thought there was *some* chance that their cajoling would work, then they are responsible and to blame for not doing so. The only case in which they wouldn't be responsible and to blame would be if they (wrongly) thought that cajoling had zero chance of working – but it's hard to see why they would have such a belief.

Making this stipulation makes the case philosophically interesting because now it is no longer clear who, if anyone (other than the attacker), bears responsibility and blame in Held's Subway. There is no bystander of whom it is true that, if she had acted differently from how she did, the attack would have been prevented. For this reason, I will from now on think about Held's Subway as containing the stipulation that any bystander cajoling alone would have been unsuccessful.<sup>7</sup>

Why might a bystander's unilateral cajoling not work? There are three possibilities. The first is a case where, even though no bystander *unilaterally* cajoling would have roused the group to action, there was a chance that one or more of the other bystanders would have cajoled simultaneously, in which case a capable group would have resulted. For this case, we need to imagine that the bystanders, upon seeing the assault, had to each quickly decide at the same time what they were going to do. They knew that if they were the only ones to cajole, their effort would have been worthless; but there was a chance that one or more of the other bystanders would also have cajoled, in which case a capable group would have been formed.

In such situations, we might be tempted to appeal to the expected utility of one's actions. We might say that bystander A, for example, is blameworthy in this case because there was a chance that, had he cajoled, he would have been the difference between a capable group forming and a capable group failing to form. We could stipulate that, given the expected utility of cajoling, the best course of action was to cajole, and that bystander A is blameworthy for

---

<sup>7</sup> This is probably also how Held conceived the case, although she didn't make it explicit in her paper. If she didn't see the case in this way, then it's hard to see why she thought her subway example produced questions of collective responsibility: the example can be explained simply by appealing to individual responsibility.

failing to do so – and the same goes for each of the bystanders B-E, by the same reasoning. This is the line of argument that Shelly Kagan has taken.

However, I will argue that Kagan's approach doesn't get things quite right. I will endorse a different approach to the problem which has been suggested by Julia Nefsky. Nefsky's approach can explain the intuition that the bystanders should have cajoled even in situations where the expected utility of cajoling was not high enough to outweigh the cost of cajoling. Drawing on Nefsky's work, I will argue the reason the bystanders are blameworthy for failing to cajole is not because of the small chance that they could have made a difference to the outcome, but because they had an obligation to help, even if they *couldn't* have made a difference.

The second and third cases where cajoling alone would be ineffectual are different, because they remove even the *possibility* that a capable group might be formed. That is, in the second and third cases, no more than one of the bystanders is prepared to come to the victim's aid.

The first of this pair of cases is a scenario in which all five of the bystanders were nasty people, who just didn't care about the welfare of their fellow passengers. Because they didn't care, they were never going to cajole, or respond to anyone else's cajoling. In this case, I will argue that the bystanders are to blame for their failure to cajole, because this action revealed an attitude of theirs which would impair the relationships that others can have with them; namely, it revealed that they were cruelly indifferent to the fate of the victim. This account of blame is due to T.M Scanlon, and I will introduce it in chapter two.

The second case is one in which there was a sole good, caring person amongst the five bystanders, and the remaining four were selfish and uncaring. Let us assume that the one good person knew for certain that, no matter what he did, none of the other five would have come to his aid – they just were not bothered by the fact that the victim was being assaulted. In this case, cajoling on the part of the good person would have been useless. It's straightforward again here that the four uncaring people bear some responsibility and blame for not cajoling. Perhaps surprisingly, I will argue that the one good person also bears some responsibility and blame in this case. He had an obligation, I will say, to perform something like a *symbolic gesture*, despite the fact that it would have had no effect. Such an act would have sufficiently disassociated him from the failure of the group to come to the victim's aid.

An important conclusion I will arrive at is that, even though I will argue that all of the bystanders (even the caring ones) bear responsibility and blame in each scenario, the reasons why they are blameworthy, and the degree of blame that they bear, will be very different in each case. Blame is not one-size-fits-all; one can be more or less blameworthy and can be blameworthy for very different reasons.

I see this thesis as having two main goals. The first is to investigate the ethical questions which arise out of Held's Subway. This goal will contain two main strands: an applied ethics strand and a theoretical ethics strand. I will be asking questions about applied ethics, and how certain individuals should have behaved in a specific (albeit fictional) situation. These will be questions such as, 'what were the bystanders' moral obligations?' and 'who is to blame for the victim's death?'. However, it will not be possible to answer these questions without the theoretical ethics strand. This strand will attempt to provide the tools required for answering

the applied ethical questions. Questions from the theoretical strand will include ‘what is responsibility?’ and ‘what is blame?’. The two strands of the thesis therefore work in tandem. As is often the case, the applied questions are what grab the attention and stir the emotions. Yet we can’t make any progress on them without a careful enquiry into the theoretical questions, which will ultimately provide us with the conceptual understanding required for the job.

The second goal of the thesis is, I think, more important. Christopher Kutz has written about what he calls the unfortunate tendency of moral philosophy to take ‘a juridical perspective, from which the disinterested writer metes out the appropriate deserts to the offending agent’.<sup>8</sup> That is not my intention in this thesis and I agree with Kutz that the tendency is unfortunate. My goal in this thesis is not to point-score or preach by identifying who is to blame in Held’s Subway. I believe that a proper understanding of how responsibility and blame apply in Held’s Subway has important implications for how we should live our lives, and for how we should respond to challenges that we face together.

In this chapter I have discussed responsibility and blame in a pre-theoretical way, motivated by what I hope are common intuitions about what those concepts are. In chapter two I conduct a careful enquiry into these concepts, in order to gain a better understanding of them, which will help to answer the harder questions thrown up by Held’s Subway. In chapter three I consider an important candidate theory about how responsibility should be apportioned in Held’s Subway: Virginia Held’s own account, according to which it is the group

---

<sup>8</sup> Kutz (2000), p.254

*itself* which is responsible. I will argue that while this approach might work well in other cases, it isn't right in the case of Held's Subway. In chapter four I will move on to discuss how the bystanders can be blameworthy for not cajoling even when doing so wouldn't have made a difference. Finally, in chapter five I will consider why all of this matters to our lives, and what it can tell us about our obligations to act together.

## **Chapter 2: Some Moral Concepts: Responsibility, Blame and Obligation**

In the previous chapter I used the words ‘responsibility’ and ‘blame’ in what I hope was a sense which most people recognise and intuitively understand. However, in order to answer the most difficult moral questions thrown up by Held’s Subway, a much more sophisticated understanding of these terms is required. This chapter seeks to pinpoint exactly which concepts we are thinking of when we ask the Responsibility and Blame Questions in Held’s Subway. I also discuss a third moral concept, obligation, which I argue we must understand in order to understand blame.

### 2.1 Responsibility

#### *2.1.1 ‘Responsibility’ and its cognates*

The words ‘responsibility’ and ‘responsible’ are put to a very wide range of uses, both in everyday language and in philosophical discussion. As an illustration, consider the following sentences, all of which contain plausible uses of one of the two words.

He is a responsible man.

Everyone is responsible for their own packed lunch.

Who is the responsible adult here?

Take responsibility for your own performance.

You are responsible for his death.

The lead surgeon is responsible for the mistakes of her trainees.

The goal of this section is to identify exactly which kind of responsibility we are interested in when we ask the Responsibility Question in Held’s Subway. To do this, it will first be necessary

to untangle the various meanings of 'responsibility' and its cognates. Only when this has been done will I be able to state clearly which kind of responsibility is of interest.

### 2.1.2 Hart's classification

In the postscript to his 1968 book 'Punishment and Responsibility', H.L.A Hart makes excellent progress on the task of classifying the various meanings of 'responsibility' and its cognates.<sup>9</sup> Rather than starting from scratch, I will sketch his classification before picking up where he left off.

Hart identifies four meanings of 'responsibility'. The first is what he calls *role-responsibility*. This is the kind of responsibility that is automatically conferred on somebody by virtue of their taking up a certain role. This understanding of the word is one way in which we use the 'responsible *for*' locution. For instance, the manager of a company is responsible for the smooth running of the company, and, possibly, for the maximisation of the company's profits. She has this responsibility by virtue of her role. She may also be given a more specific task – perhaps to manage a project which must deliver a report in three weeks' time. In this instance, her role as leader of the project means that she is responsible for the successful running of the project. There are many well-known roles in society which confer such role-responsibility – consider: babysitter, councillor, librarian, teacher, etc.

Note also that the responsibility-conferring role need not be a formal position. If you and I plan to have a picnic, and we agree that I will get the sandwiches and you will get the drinks,

---

<sup>9</sup> Hart (1968), p.210-230

I take on the role of Person In Charge of Bringing the Sandwiches, and you take on the role of Person In Charge of Bringing the Drinks, and those roles confer on us responsibility. The notion of role-responsibility also gives rise to the locution ‘my responsibilities’; the Person In Charge of Bringing the Sandwiches has certain responsibilities – to go to the shop in advance, to wrap the sandwiches correctly, etc.

Another locution relating to role-responsibility is ‘she is a responsible person’. Usually, this means something along the lines of ‘you can trust her to adequately fulfil her role-responsibilities; she will make serious efforts to fulfil her responsibilities’, and so on. Sometimes, though, this locution can also be used to refer to capacity-responsibility, which I discuss below.

The second meaning of ‘responsibility’ that Hart talks about is *causal responsibility*. Causal responsibility has to do with identifying the perpetrator of the physical action which caused some event; what Peter French has called responsibility in the ‘who-dun-it or what-dun-it’ sense.<sup>10</sup> Often, it’s very straightforward to assign causal responsibility. When your 18-month-old child scribbles crayon all over the wall, she is causally responsible for the scribbles. When the dog goes to the toilet on the living room floor, the dog is causally responsible for the mess.

Sometimes, though, it’s less easy to establish the agent who is causally responsible for some event. 14-year-olds in the UK are often taught that the assassination of Archduke Franz Ferdinand caused the First World War. Does this mean that his assassin is causally responsible

---

<sup>10</sup> French (1979), p.210

for the First World War? Such questions are difficult and beyond my scope here, but it does seem to be stretching the concept of causal responsibility to go along with such statements. After all, there were presumably ample chances for war to be averted after the Archduke's death.

Nevertheless, for my purposes, the point stands that causal responsibility is a commonly recognised, commonly-assigned kind of responsibility, and must be kept conceptually separate from role-responsibility and all the other kinds of responsibility which are to follow.

Third, Hart alerts us to the existence of *legal liability-responsibility*. This is responsibility in the sense of being liable to be made to face the consequences of some act. It's important to note that causal responsibility is neither necessary nor sufficient for legal liability-responsibility. For instance, if you hire a builder to change the windows in your house, and you sign a contract with the building company which stipulates that the builder can't be sued in the event of any damage to your property, then if the builder drops a window and it smashes, he is causally responsible but not legal liability-responsible for the smashed window. His causal responsibility is not sufficient for legal liability-responsibility.

On the other hand, if a 16-year-old borrows his mother's car and crashes it, resulting in a large insurance bill, the mother would be legal liability-responsible for paying the bill, since the child is not yet an adult (and, let us imagine, has no money with which to pay the bill anyway). So causal responsibility is also not necessary for legal liability-responsibility.

Legal liability-responsibility is an interesting phenomenon in its own right. In tort law, often what the courts are most concerned with is not who is culpable or blameworthy for some crime, but who is going to make it up to the victim.<sup>11</sup> It's about finding someone to make amends – often this turns out to be the person with the deepest pockets, for instance in cases where employees of a large company cause some kind of damage, and the company itself must pay the cost.<sup>12</sup>

The concept of legal liability-responsibility can be extended to non-legal settings as well. If I borrow a friend's pen and lose it, it's plausible that I am responsible for replacing it, or at least making up for it in some other way.

Finally, Hart talks about *capacity-responsibility*. To say that someone is capacity-responsible is to assert that they have, to a sufficient degree, the mental capacities required to understand legal and moral rules, to reason correctly, to make decisions, and to generally be in control of their conduct. When we say of someone that they are responsible for their actions, we mean that they are capacity-responsible. As Hart points out, someone who lacked some or all of these faculties may be said to be 'suffering from diminished responsibility'.<sup>13</sup> As mentioned above, sometimes when we utter sentences such as 'she is a responsible person', we refer to a subject's capacity-responsibility. On other occasions we may be referring to their role-responsibility.

---

<sup>11</sup> See Rosen (2004), p.303-304

<sup>12</sup> Joel Feinberg makes this point in (1968), p.676.

<sup>13</sup> Hart (1968), p.228

Role-responsibility, causal responsibility, legal liability-responsibility, capacity-responsibility: this concludes Hart's classification of the word 'responsibility' and its cognates. However, when I ask myself who is responsible in Held's Subway, I find that none of these four concepts exactly specifies the question that most interests me. There certainly are important questions about, for instance, who is or should be legal liability-responsible for the death. Perhaps also there are questions of what role-responsibility was in play – perhaps each bystander failed to live up to their role-responsibility which is conferred by their 'role' as the victim's fellow citizens. There is also of course a causal question, although it's relatively easy to answer: the attacker is causally responsible for the death.

What interests me most, though, is something altogether different. What I am interested in is that most elusive of concepts, *moral responsibility*.

### 2.1.3 Moral responsibility

Hart does not list moral responsibility as one of the meanings of 'responsibility' he is interested in.<sup>14</sup> However, he does mention it briefly at two points. First, he provides an argument that legal liability-responsibility and moral blameworthiness need not coincide.<sup>15</sup> That is, one can be legal liability-responsible yet be in no way morally blameworthy. Second, he speculates that the account he gives of legal liability-responsibility can be transcribed *mutatis mutandis* onto moral responsibility.<sup>16</sup> The resulting conception of moral responsibility would then be that 'to say that a person is morally responsible for something he has done or

---

<sup>14</sup> Ibid. p.212

<sup>15</sup> Ibid. p. 222-223

<sup>16</sup> Ibid. p.225-277

for some harmful outcome of his own or others' conduct, is to say that he is morally blameworthy, or morally obliged to make amends for the harm'.<sup>17</sup>

These remarks aside, Hart does not spend much time considering moral responsibility, and does not go into anywhere near the level of detail which he does for the other meanings of 'responsibility'. Thus, in order to get a better idea of the kind of responsibility I am interested in, I will have to draw on other resources. Hart's work is helpful for distinguishing various kinds of responsibility, but now what is required is a careful enquiry into the nature of moral responsibility.<sup>18</sup>

#### *Scanlon's view*

What, then, is the best account of moral responsibility? What does it mean to say that someone was morally responsible for something they did? I think that the best account is given by T.M. Scanlon. He writes the following.

*Questions of "moral responsibility" are most often questions about whether some action can be attributed to an agent in the way that is required in order for it to be a basis for moral appraisal. I will call this sense of responsibility responsibility as attributability. To say that a person is responsible, in this sense, for a given action*

---

<sup>17</sup> Ibid. p.225

<sup>18</sup> For the purposes of this thesis, I will be assuming that moral responsibility is a real phenomenon which can sometimes be correctly attributed. That is, I will be ignoring arguments that say we should be sceptics about moral responsibility, or that we can never be justified in ascribing moral responsibility. For an example of such an argument, see Rosen (2004).

*is only to say that it is appropriate to take it as a basis of moral appraisal of that person.*<sup>19</sup>

In other words, you are morally responsible for your action simply if it is appropriate to include the fact that you performed that action in a moral appraisal of you as a person. There are three interesting features of this view which I would like to highlight.

Firstly, on this account of moral responsibility, moral responsibility is not a gradable concept. That is, it doesn't make sense to say that somebody is morally responsible to a greater or lesser degree for some action. Moral responsibility is all-or-nothing: either it is appropriate to include the fact that you performed some action in a moral appraisal of your character, or it is not appropriate. I think that this goes against the pre-philosophical understanding of moral responsibility and, indeed, against the view of many philosophers who say that moral responsibility is a gradable concept.<sup>20</sup> Commonplace utterances such as 'I may be morally responsible for that act, but I am less morally responsible than you' are mistaken on Scanlon's view. Scanlon himself does not make this explicit, but I think that it is an undeniable (and in no way damaging) aspect of his account.

Secondly, on Scanlon's account the question of moral responsibility is conceptually prior to questions about praise, blame, reward or punishment. As Scanlon writes regarding the moral appraisal in question, 'nothing is implied about what this appraisal should be – that is to say,

---

<sup>19</sup> Scanlon (1998), p.248, his emphasis.

<sup>20</sup> See, for example, May (1992), p.120: 'Moral responsibility is a scalar concept; there are degrees of responsibility'.

about whether the action is praiseworthy, blameworthy, or morally indifferent'.<sup>21</sup> We first must decide whether or not someone's action can be used as a basis for moral appraisal, and only if we answer in the affirmative can we then move on to ask questions about what this appraisal will be. This is an important feature of the view because it acknowledges a strict but often neglected distinction between moral responsibility and blame. Often we use the two terms interchangeably, a verdict of moral responsibility being taken as equivalent to a verdict of blame. We must resist this practice. That I am morally responsible for an act will just as often make me worthy of praise as of blame.

The question of praise and blame is where gradability fits in to the view. While on Scanlon's view we can't say that someone is more morally responsible than another, we can certainly say that someone is more praiseworthy, or more blameworthy, than another. Praise and blame naturally come in degrees. Suppose that you and I have a mutual friend who has fallen on some hard times. I give him £5 and consider my job done. You give him £1000 and also help him to get his spending in order and apply for new jobs, and you let him sleep in your flat whenever he wants to.<sup>22</sup> We are both equally morally responsible for our actions, in the sense that both of our actions can legitimately be used to appraise our respective moral characters. Yet it is clear that you are much more praiseworthy than I am.

The final point to note about Scanlon's account is that a verdict of moral responsibility in no way entails that the agent is required to make amends for anything.<sup>23</sup> In this way, Scanlon's

---

<sup>21</sup> Scanlon (1998), p.248

<sup>22</sup> For this example, suppose that you and I have roughly equal means so that £1 is worth the same to both of us.

<sup>23</sup> See Collins (2017) for an important discussion on how our past misdeeds can confer duties on us to make amends.

account differs from Hart's speculative account which I mentioned earlier: for Hart, being morally responsible for something entails that you are 'morally obliged to make amends for the harm'. Notice that Hart here also makes the assumption I just referred to – that moral responsibility is something negative. For Scanlon, though, the question of what people are required to do or not do for one another, which he calls *substantive responsibility*, is entirely separate to the question of moral responsibility.

I take this to be a strength of Scanlon's account. The failure to keep substantive responsibility and responsibility as attributability separate leads to the commonly-heard argument that because someone is morally responsible for something in the attributive sense, it must follow that they owe something to somebody, or that they deserve no help in improving their situation. It allows people to be written off, and permits the idea that if people are morally responsible for their actions then they must live with the consequences alone, because it is their own fault.<sup>24</sup>

To illustrate why the two senses of responsibility should be kept strictly apart, Scanlon uses the example of the 'willing addict'. Imagine a drug addict who grew up in a part of town where drugs were freely available and taking them was what young men did. Even though he knew of the dangers, he willingly got into this lifestyle, under no direct coercion from his peers. Over time he became addicted, and he never made a concerted effort to prevent this from happening. Suppose now, though, he has children and a mortgage to pay, which he struggles to do because he spends all his money on his addiction. Those who do not recognise the

---

<sup>24</sup> See Scanlon (1992), p.293

distinction Scanlon draws may say that it is the addict's responsibility alone to care for his children and pay his bills. Using Scanlon's account, though, we are able to say that even though he voluntarily got into drugs in the first place, this does not absolve the state from the responsibility of helping him and his family to get by. The normalcy of taking drugs in the area he grew up in, and the difficulty of overcoming addiction once it has taken hold mean that he can't be expected to bear full substantive responsibility for his condition. He may be responsible attributively, but he does not bear full substantive responsibility for ameliorating his situation.<sup>25</sup>

I think Scanlon's account of moral responsibility is right. There are now just a few more distinctions to be drawn and clarifications to be made before concluding this enquiry into moral responsibility.

#### *Forward-looking and backward-looking responsibility*

Some philosophers have drawn a distinction between forward- and backward-looking responsibility. Thus Carol Rovane writes 'we should construe forward-looking responsibility not as a matter of *bearing* responsibility in the future for our present (or past) deeds, but rather as a matter of *taking* responsibility for what will happen in the future, by doing things in the present with an eye to what their future consequences will be'.<sup>26</sup>

Rovane is surely right that this is a way we use the word 'responsibility'. We often encourage people to think through the probable consequences of their actions, and ask them whether

---

<sup>25</sup> Ibid. p.292

<sup>26</sup> Rovane (2014), p.12, her emphasis.

they want to be responsible for those consequences. This is an important facet of our moral education, and is a useful tool for ensuring that bad things don't happen. However, I mention this kind of forward-looking responsibility only to make the point that this is not the kind of responsibility I am interested in in Held's Subway. I am taking the standpoint of somebody being told about an event, Held's Subway, that has already happened, and thinking about who is morally responsible for what they did or didn't do. Therefore, my interest is in backward-looking responsibility; trying to ascertain whose actions were suitable bases for moral appraisal of their character. Forward-looking responsibility, on Rovane's account, doesn't qualify as moral responsibility in the way I am understanding moral responsibility.

*On the distinctness of moral responsibility from other kinds of responsibility*

Peter French has written that 'moral responsibility...is not a class apart but an extension of ordinary, garden-variety, responsibility'.<sup>27</sup> Based on his comments which precede this sentence, I take him to mean that, just as in the case of 'ordinary, garden-variety responsibility', moral responsibility depends on the existence of a 'responsibility relationship'. The difference in the moral responsibility case is that no formal relationship needs to be established, because we are all, by default, in responsibility relationships with our fellow humans when it comes to moral responsibility.

This may be true on French's understanding of moral responsibility. However, on my understanding, moral responsibility must be seen as a class apart from the other forms of responsibility which Hart identified. This is because it is possible for a person to bear all of

---

<sup>27</sup> French (1979), p.211

these other forms of responsibility without being also morally responsible. To demonstrate this, I will give four examples.

1. A manager working for a large corporation is given the task of delivering a report for a client by 5pm on Friday afternoon. She is given a team of three junior employees to work on the report. Everything is going well throughout the week, and the report is on schedule. Then on Thursday afternoon, all three members of her team quit their jobs unexpectedly. None of the employees' decisions to quit were due to the manager, and there was no way in which they could have been foreseen. As a result of the departures, the manager doesn't complete the report in time, and the deadline is missed. In this scenario the manager is *role-responsible* but not *morally responsible* for the failure of the project. She is not morally responsible because her action (failing to deliver the report on time) is not an appropriate basis of a moral appraisal of her character. It would be wrong to conclude, from the fact that the report wasn't finished on time, that she is morally deficient in some way.

As an aside, note that if the report wasn't completed because the manager didn't work hard enough, or didn't plan sufficiently, or due to some other incompetence, this may make the manager morally responsible. A failure to adequately plan the project, for instance, might indicate a lack of commitment or diligence in her work, or a lack of concern for the client's needs, or her duties to her employer.

2. A two-year-old gets hold of some crayons and draws all over the wallpaper of his parents' living room. The child is *casually responsible* but not *morally responsible*. We don't generally tend to take *any* actions of a young child to be bases for a moral appraisal of that child. To do

so would be to ignore the fact that children must first be educated in how to act morally, and until they have been, must not be judged for their failure to do so.

3. A 16-year-old boy borrows his mother's car and drives it very carefully. However, through sheer bad luck, he swerves on a patch of ice and crashes into a car parked by the side of the road, running up a large insurance bill. When the bill comes, the mother must pay, since her son is not yet an adult, and anyway has no money of his own. The mother is *legal liability-responsible* but not *morally responsible* for the damage. We shouldn't take the fact that her son crashed the car as a basis for a moral appraisal of his mother.

Note that, if the crash was due to reckless driving on the son's part, the mother may bear some responsibility, as we might take her to have neglected to sufficiently educate her son on the importance of safe driving.

4. A fully-cognitively-functioning 32-year-old man is standing up on a train journey. The train stops abruptly, causing him to involuntarily take a step backwards, which results in him stepping on the toe of one of his fellow passengers. The man is *capacity-responsible*, but he is not *morally responsible* for stepping on the toe. We shouldn't take that action as a basis for a moral appraisal of his character.

What I take these examples to show is that moral responsibility as I understand it is importantly distinct from other kinds of responsibility. Being responsible in other ways doesn't entail that one is also morally responsible.

My goal in this section was to identify exactly what kind of responsibility I am interested in when I ask the Responsibility Question in Held's Subway. I found that none of the kinds of responsibility identified by Hart adequately described what I was interested in; rather, I am interested in *backward-looking moral responsibility*. I take the interesting question to be, following Scanlon, whether someone's actions can be plausibly used as a basis of a moral appraisal of their character.

## 2.2 Blame

### *2.2.1 Why ask the Blame Question?*

As I have mentioned, the Responsibility and Blame Questions are often conflated, especially in our everyday talk. People often attribute moral responsibility to someone as a way of expressing their belief that they are to blame. One might say, for instance, 'you are morally responsible for his death' in an accusatory tone of voice. I suspect that the amalgamation of these two different questions into one is a result of an excessive focus on moral responsibility as a negative concept. For whatever reason, our everyday usage of the term brings connotations of guilt or wrongdoing; we rarely use the term to convey positive appraisal of one's actions such as bravery or competence.

An important difference between moral responsibility and blame, as I am understanding them, is that the latter is gradable whereas the former is not. When questions about moral responsibility are taken to be questions about whether a particular action can be appropriately taken as a basis for the moral appraisal of one's character, it makes no sense to ask whether one is more or less morally responsible than somebody else. Either the action is an appropriate basis or it isn't. Blame, on the other hand, comes in degrees. As my analysis of

blame will show, it is perfectly coherent to say that someone is blameworthy to a greater or lesser extent than somebody else.

Before beginning my enquiry into what blame is, and who is blameworthy in Held's Subway, I will pause to ask whether we should even be asking this question, and why it matters. Blame often gets a bad press; seeking to blame somebody is seen as petty or vindictive, or a way of keeping score on people's moral record. Michael Zimmerman, for instance, writes that 'I take [moral blame] to be a sort of censure, from the moral point of view, of one's character as manifested in the production of the outcome in question – a censure of the sort, one might say, which amounts to inscribing a "demerit" on the "report-card of life"'.<sup>28</sup> This understanding of blame could justifiably be viewed as a kind of self-righteous bookkeeping, serving only to separate the morally good from the morally bad, and condemn those who fall short of certain standards. It suggests a desire to preach, or simply to make one feel better about oneself.

Blame needn't be like this. Miranda Fricker has recently argued persuasively that blame can have a much more positive use. Fricker argues that in the paradigm case of blaming, in which 'I wrong you, and in response you let me know with feeling that I am at fault for it',<sup>29</sup> the act of blaming can help to bring 'increased alignment between the moral understandings of blamer and wrongdoer'.<sup>30</sup> The thought is that blaming can be a tutelary experience for both blamer and wrongdoer: the blamer lets the wrongdoer know that she feels he has done wrong, he perhaps gives a defence and explanation of his actions, and the dialogue goes back

---

<sup>28</sup> Zimmerman (1985), p.115. Zimmerman is quoting from Downie (1969).

<sup>29</sup> Fricker (2016), p.171

<sup>30</sup> Ibid. p.173

and forth, with the (ideal) result that the two people become closer to one another in their moral beliefs. Fricker also suggests that blame can be proleptic: treating someone in the way that you would treat a morally thoughtful, sensitive person can increase the likelihood that they will become such a person. Seen in this way, blame can serve two very important purposes: it can teach and improve both blamer and blamed, and it can improve our relationships with others by expressing the expectations we have of their conduct.

This is the understanding of the purpose of blame that I will have in mind during my enquiry. My hope is that understanding who is to blame, why, and to what extent in Held's Subway will provide an insight into how we should behave in real-world cases which resemble Held's Subway in certain ways. In short, I think that we should ask the blame question in Held's Subway because it can teach us something about how we should live our lives.

### *2.2.2 Scanlon's account of blame*

The account of blame I favour is the one given by Scanlon in his 2008 book 'Moral Dimensions'. Scanlon argues that to say that a person is blameworthy for an action is to say that the action reveals 'something about the agent's attitudes toward others that impairs the relations that others can have with him or her'.<sup>31</sup> I blame you for an action if I judge you to be blameworthy and I take our relationship to be 'modified in a way that this judgment of impaired relations hold to be appropriate'<sup>32</sup>. Note the difference between blameworthiness and blaming: I can judge that somebody is blameworthy even though I do not actually blame them. I might

---

<sup>31</sup> Scanlon (2008), p.128

<sup>32</sup> Ibid. p.128

acknowledge that somebody's conduct gives me reason to hold the attitude appropriate to a new, impaired, relationship, yet I may, for whatever reason, fail to actually hold that attitude.

Scanlon gives the example of a friend, Joe, who, at a party which you are not attending, makes several unpleasant jokes about you and laughs along when others do the same.<sup>33</sup> Blaming Joe for this behaviour would involve arriving at a judgement that Joe's actions reveal a certain attitude he has towards you and your friendship, and to revise your attitude to Joe accordingly. You might start to think that you shouldn't tell Joe your intimate secrets, for instance, because he is likely to reveal them to others in order to make friends at future parties.

Following Fricker's account, the most productive way of blaming Joe would be to convey these thoughts to him, so that he will realise your reaction to his behaviour at the party. He may then come to see that he has failed in some way to uphold the standards required in a friendship. Ideally, he will apologise and promise to change his ways, and to make up for it. Your blaming of Joe need not be irreversible; if he mends his ways and adequately apologises, it is open to you to reverse your position, and come to see your relationship with Joe as it was before the party incident.

One aspect of Scanlon's account of blame that merits attention is that blame is clearly dependent on prior expectations of the wrongdoer's behaviour. The reason you might blame Joe is because he has failed to live up to certain *obligations* of friendship: specifically, the

---

<sup>33</sup> Ibid. p.129

obligation not to make cruel jokes at a friend's expense in order to impress people at a party. Suppose that Joe had not been a friend but merely someone you vaguely knew – a distant colleague, for instance, or somebody you studied with at school but haven't seen for twenty years. If you found out that such a person had been badmouthing you at parties, you would, of course, feel hurt. However, the keenness of the betrayal wouldn't be so acute; while you might conceivably blame the person, it wouldn't be the same as if Joe, your close friend, had done it. The way you took your relationship, such as it was, with that person to be impaired would be less severe than in the case of Joe. Suppose now that it was your sworn enemy telling cruel jokes about you – in this case, you might not take your 'relationship' to be impaired at all.

The key role which obligations play in blame can be seen through another example. If I forget my mother's birthday and don't send her a card or even phone her, she has good reason to blame me. She may take my action to reveal a cold indifference on the part of her son towards his mother, and may take our relationship to be, at least temporarily, impaired. However, if my friend Paul fails to send a card to or phone up my mother on her birthday, my mother won't consider her relationship with Paul to be in any way impaired. There is no obligation on Paul to remember my mother's birthday, whereas there is such an obligation on me.

It is apparent, based on these considerations, that answering the blame question is going to require a more thorough understanding of moral obligation.

## 2.3 Obligation

My finding from the previous section was that on Scanlon's account of blame, whereby to blame somebody for an action is to hold that your relationship with them is impaired in some way as a result of the attitude which their action has revealed in them, relies on an understanding of obligation. We can't blame people unless we knew what was required of them in a certain situation. This section, then, will be devoted to an enquiry about moral obligations, and how they might arise. Specifically, I will be asking whether moral obligations can arise simply as a result of certain capabilities which people have.

### *2.3.1 When does 'can' imply 'ought'?*

#### 2.3.1.1 The Individual Case

It is commonly accepted in moral philosophy that 'ought' implies 'can'. That is, if somebody morally ought to do something, it follows that they can do that thing. In other words, nobody is morally required to do something that they are unable to do. To place moral obligations on people which they are unable to fulfil would make morality overly demanding, and lead to widespread failure to fulfil our moral obligations. In this section I want to enquire as to whether the inverse of 'ought' implies 'can' is ever true. That is, is it ever the case that 'can' implies 'ought'?

There are certain cases where it seems obviously right that 'can' implies 'ought'. If I am passing by a small, shallow pond and see a child drowning, the fact that I *can* easily bend down

and lift the child out seems to confer on me an *obligation* to do so. If I just walk on by, I have flouted my moral obligation.<sup>34</sup>

This, of course, is Peter Singer's example. Singer writes that 'a plausible principle that would support the judgement that I ought to pull the child out is this: if it is in our power to prevent something very bad happening, without thereby sacrificing anything of comparable significance, we ought to do it'.<sup>35</sup> Here, Singer is suggesting that, in this particular scenario, 'can' implies 'ought'. Singer also notes that it is not just consequentialists who would accept such a principle; most non-consequentialists also believe that we should prevent bad outcomes when nothing else of moral significance is at stake. In the pond case, there are no rights being violated, or categorical imperatives flouted, if I bend down to lift the child out of the pond.

On the other hand, there are cases where 'can' doesn't seem to imply 'ought' at all. Right now, I am able, if I chose to, to run naked through Russell Square. But there is no sense in which I morally ought to do so.

The question, then, is: when does my ability to do something confer on me a moral responsibility to do that thing? Singer suggests two conditions: the act must prevent something very bad from happening; and it must not involve any sacrifice of comparable significance. We might go further than this though. If it is in my power to do an act which will

---

<sup>34</sup> It may be objected that if saving the child conflicts with another activity which I can and should do, for instance saving ten children from drowning in a pond further along the road, this overrides my obligation to save the single child. For simplicity, I will ignore such cases and focus only on cases where there is no outside option to perform some act which might be morally better.

<sup>35</sup> Singer (2011), p.199

ensure that something *good* happens, without significant sacrifice, I ought to do that, too. If I can easily get my friend, a huge Rolling Stones fan, free tickets to their concert through a friend who works for the band, it is plausible that I ought to do so. Not to do so could be seen as failing in my moral duty as a friend, since I had an easy opportunity to do something which would have significantly increased my friend's happiness at no cost to me.

A hard-line Act Utilitarian would have a straightforward answer to the question of when 'can' implies 'ought'. Simply, if you are able to perform act A, and A will bring about the state of affairs with maximum utility, you have an obligation to do A. This answer seems excessively demanding: it means that, if I have to choose between saving my brother and saving three strangers, I should save the strangers. Intuitively, it seems at least *permissible* in this case for me to save my brother. That is, saying that I am morally required to save the strangers ignores the morally significant personal cost I will have to bear in doing so.

I propose instead that we accept the following.

'Can' implies 'ought' (Individual)

That S *can* do A implies that S has an *obligation* to attempt to do A when:

1. If A-ing imposes a cost on S, that cost is sufficiently small for S.
2. The potential upside of doing A is sufficiently high.
3. S knows that S can do A, and also knows both 1 and 2.

Several remarks are needed. First, note that *S*'s obligation is to *attempt* to do A, not to do A. There may be many factors out of *S*'s control which mean that the attempt is unsuccessful. Going back to Singer's pond example; perhaps I make a genuine attempt to save the child, and my physical capabilities are such that such an act is achievable for me. However, on this occasion, there happens to be an unusually slippery stone on the bed of the pond and I step on it, losing my footing and missing my chance to save the child. I suggest that, in this example, I am not morally blameworthy. My actions can't reasonably be construed as evidence of a bad moral attitude of mine, such that somebody (the child's mother, for instance) could see my act as a reason to modify their relationship with me. At most I can be accused of clumsiness, although given that I couldn't have known about the stone, even this seems harsh.

Second, note that condition 1 makes reference to the cost for *S*, rather than the cost overall. This condition is what differentiates this account from a straightforward Act Utilitarian account. If A-ing would impose to great a cost on the agent, the agent no longer has an obligation to A. This is what allows me to say that it is permissible for me to save my brother in the case where I must choose between my brother and three strangers – the moral cost to me of failing to save my brother is great enough to override what would otherwise be an obligation to save the three. Note that condition 1 also makes allowances for the fact that an objectively equal cost, borne by two different people, may be very different from the point of view of the subjects. For instance, suppose that a man needs £100 to be able to afford a live-saving medical procedure. Assume that the procedure is guaranteed to save his life, his undergoing the procedure isn't going to preclude anybody else from having it, and nobody else will be harmed as

a result of the medical professional taking time out of her day to perform the procedure. If this man were standing on the street with a sign asking somebody for £100, and a billionaire walks by with several hundred pounds in cash in his wallet, I suggest that the fact that the billionaire *can* provide the £100, coupled with the fact that this is a negligible amount of money for him, means that he has *an obligation* to provide the £100. On the other hand, if a very poor person, for whom £100 is a much more significant sum of money, walked past the man, there would be no obligation to hand it over. The same objective cost can be very different for different agents.

A final note about condition 1 is that what qualifies as a 'sufficiently small' cost will vary according to the size of the upside. Generally speaking, agents must be expected to undergo greater costs as the expected upside gets greater, other things being equal.

Condition 2 refers to the *potential* upside of A-ing. This acknowledges that expected outcomes are important: it doesn't have to be guaranteed that A-ing will bring about the good outcome for the agent to have an obligation to A. For instance, in the pond case, I needn't be certain that my attempt to save the child will be successful. Since the potential upside of my doing so is very large (saving a child's life), it is enough that there is some chance that I will be successful. In referring to upsides, condition 2 also acknowledges that the outcomes can be either good or bad; both can confer obligations on agents.

Condition 3 is the epistemic condition for 'can' implies 'ought'. It requires that S know that she is able to do A, that A-ing would not impose an excessive cost on her, and that

the potential upside of A-ing is sufficiently large. Note that condition 3 does not require that S know exactly *how* to A; all that is required is that she know that A-ing is within her capabilities. Michael Doan has persuasively argued that it is not necessary that the agent know exactly how to perform the task. Rather, so long as the agent knows that she *can* A, she has an obligation to conduct an enquiry into how to A.<sup>36</sup> This obligation to enquire can be absorbed in my use of the word ‘attempt’: conducting a genuine enquiry into how to A can be seen as attempting to A when one doesn’t yet know how to A.

### 2.3.1.2 The Collective Case

In the previous section I tried to get an understanding of when ‘can’ implies ‘ought’ in the case of individual actors. Given that Held’s Subway, the object of enquiry in this thesis, concerns group inaction, I will attempt in this section to develop this account for collective cases.

The first thing to note is that there intuitively seem to be cases where ‘can’ implies ‘ought’ for groups in the sense that there is something which the group can do together which confers obligations on the individual members. For instance, suppose that a company has the chance to make a deal which will ensure that a planned programme of redundancies can be cancelled. It is plausible that the company has an obligation to try very hard to make this deal (supposing that there is no significant sacrifice involved in making the deal). This means that there are certain obligations on some or all of the members of the company. What are the obligations which the members bear?

---

<sup>36</sup> See Doan (2016), especially section 4. Doan is talking about groups of people engaging in what he calls ‘collective inquiry’. I have adapted his discussion for the individual case here.

This kind of case is the one with which we are faced in Held's Subway. The group of bystanders has the ability to intervene in the attack and save the victim. This appears to impose certain obligations on the members of the group. What obligations are these?

One might be tempted by the following.

'Can' implies 'ought' (Collective)

That a group G *can* do A implies each member of G has an *obligation* to attempt to do her part in G A-ing when:

1. If doing her part in the group's A-ing imposes a cost on a member, the cost is sufficiently small for that member.
2. The potential upside of doing A is sufficiently high.
3. Each member knows that G can do A, and also knows both 1 and 2.

This is an initially plausible solution. In the case of the business, it means that each member of the company is obliged to do whatever their part is in making the deal happen. This might mean, for example, that the CEO has to travel to a client's offices, the junior employees have to write up the correct contracts, and so on.

However, this kind of response, whereby the obligation borne by each member of a group is to *do one's part* in a collective action faces serious problems. This is shown by examples of the following kind, which Alexander Dietz calls *Gibbard cases*.<sup>37</sup>

Consider a two-player interaction game with the following payoffs.

|   |      | You          |               |
|---|------|--------------|---------------|
|   |      | Do A         | Do B          |
| I | Do A | Second best  | Worst for you |
|   | Do B | Worst for me | Best          |

In this scenario, there is a clear sense in which we should both do B. That is, we should together perform the joint action (B,B), since this will bring about the best results. The second best outcome would be if we both did A, whereas if I do B and you do A, that gives me my worst possible outcome, and vice versa. This type of game has often been characterised as a Stag Hunt: if we both hunt a stag (B), we will catch one, which is the best possible outcome. If we both hunt a hare (A), we will catch one, although this is not as good as a stag. However, if I hunt a stag and you hunt a hare, I will end up with nothing and you will end up with a hare, and vice versa.

---

<sup>37</sup> Dietz (2016) p.960. Dietz takes the case from Allan Gibbard – see Gibbard (1965), p.214-215. The example is also discussed by Derek Parfit in (1984), p.72.

This kind of example shows the problem with the requirement to unilaterally do one's role in a group action. According to the account I just gave, it would be my obligation in this scenario to attempt to do B, because that would constitute doing my part in (B,B), which is the best action that we together can do. However, it's easy to see that if I do my part, but you defect and do A, we will end up in the bottom-left quadrant of the outcome space: the joint-worst possible outcome for me. Given what you are doing, it's best for me *not* to attempt to do my part in our best possible action – doing so would bring about my worst outcome.

There are other cases in which unilaterally doing my part in a collective action simply doesn't make sense.<sup>38</sup> Suppose that I am a member of a circus troop, and my role is to be on the top of the human pyramid during the finale of the show. Suppose that the rest of the troop all go on strike one evening to protest about poor working conditions in the circus. When the finale comes, according to the idea under consideration, I am obliged to attempt to do my part in the group action, meaning that I must try to take up my position atop a non-existent human pyramid. Such examples show that unilateral execution of one's part in a joint action can't be what is morally required of group members.

Indeed, in Held's Subway we have seen that such unilateral performance of one's part won't work. Any members of the random collection who unilaterally intervene in the attack will be unsuccessful, and are likely to cause themselves some harm in the process. This is why, as I mentioned in chapter 1, the bystanders can't be held responsible or blameworthy for not unilaterally intervening.

---

<sup>38</sup> Thanks to Han van Wietmarschen for pointing this out.

Dietz offers a way out of this puzzle. The injunction to do one's part no matter what, which he calls The Simple View, must be amended in order to make it plausible. The first alteration he suggests is that we must focus only on cooperative cases. In Gibbard's case, as Felix Pinkert has pointed out in his discussion of these cases, we 'cannot come together to form a common strategy'.<sup>39</sup> This leaves us open to achieving the worst outcome if one of us does B and the other defects to A. Dietz's solution is that we should restrict our analysis to cases where we are acting cooperatively. He endorses what he calls Jackson's Principle.

***Jackson's Principle:*** *If we together ought to act in some way, and we will act in this way if I do my part, then I ought to do my part.*<sup>40</sup>

That is, if everything is in place so that all that has to happen in order for us to do some good act is that I do my part, I am required to do my part. In Gibbard's case, this means that I am required to do B only if I know for sure that you are going to do B.<sup>41</sup>

The trouble with this view, as I see it, is that it creates too strong a segregation of cooperative and non-cooperative cases. It appears to avoid a difficult question: what should we do in non-cooperative cases, where I cannot be sure that the others will do their part? Dietz ignores this

---

<sup>39</sup> Pinkert (2015), p.973

<sup>40</sup> Dietz (2016), p.973. Dietz is paraphrasing a principle given by Frank Jackson in (1987), p.107.

<sup>41</sup> Dietz offers two further amendments, but they are not pertinent to my discussion here. Briefly, they are that Dietz couches his position in terms of having a *reason* to do one's part, to allow for the possibility that there may be other requirements of you, such as saving your child, which override your obligation to do your part. Secondly, he specifies that our group-based reasons and our individual reasons must not be double-counted.

problem, asking us instead to focus on co-operative cases where I *can* be sure that the others will do their part.

Can we do any better than this? It would be useful to have an answer to the question of what to do in the unfortunate situation where everything is not as easy as we would like it to be, and we cannot be sure of our fellow group members' contributions. I will attempt to give such a theory now.

I propose the following.

'Can' implies 'ought' (Collective) – revised

That a group G *can* do A implies each member of G has an *obligation* to attempt to do her part in ensuring that the group members coordinate in the way required for A-ing when:

1. If doing her part in ensuring that the group members coordinate in the required way imposes a cost on a member, the cost is sufficiently small for that member.
2. The potential upside of doing A is sufficiently high.
3. Each member knows that G can do A, and also knows both 1 and 2.

Note the difference between the revised version and the version I considered before: the obligation in the revised version is not to attempt to do one's part in the group's A-ing, but to

attempt to do one's part in *ensuring that the group members coordinate in the way which Acting requires*.

Again, note that the obligation is only to attempt. For myriad reasons outside of our control, our best attempts may sometimes fail. If I attempt to do my part in ensuring that the members coordinate, but am struck down by illness at the crucial moment, it is unreasonable to accuse me of a moral failure.

The obligation to ensure that the group members coordinate in the desired way will mean different things depending on the kind of group in question. In highly organised, hierarchical organisations such as Google, rank-and-file workers may find that they can fulfil their obligations simply by carrying out their usual duties, and refraining from actively working against the functioning of the organisation. This is perhaps as much as they can do to ensure that the group coordinates in the right way. Executives, on the other hand, may have to do more to fulfil their obligations: they may have to provide instructions, give talks to staff, speak to the press etc. in order to make sure that the organisation coordinates in the right way.

What of Held's Subway? The obligation to attempt to ensure that the members coordinate in the right way means that each member is obliged to attempt to bring the group together into a unit which will be able to intervene in the attack. This obligation is derived from the potential that the group has to stop the attack. This is the obligation which led to the conclusion from chapter 1 that what was required of the bystanders in Held's Subway was to cajole (so long as there was a chance that unilateral cajoling might work – we have not yet arrived at the difficult case where the bystanders know that cajoling alone will be ineffective). The fact that, as Held

presents the case, none of them cajoled means that they have all flouted a moral obligation – they failed to do something that was morally required of them.

How might this failure result in blame? Recall that on Scanlon's view of blame one is blameworthy for an action if that action reveals an attitude on the part of the actor which indicates that one's relationship with that person is impaired in some way and in need of modification. I propose that the bystanders in Held's Subway have flouted a moral obligation, and this indicates a morally deficient attitude. It reveals a lack of concern for the victim, their fellow passenger, and a coldness which would make us think twice about entering into certain relationships with them. If one were to meet one of the bystanders after the event, one would be unlikely to trust him with any task requiring the care of another person, because he has just sat by and witnessed somebody being beaten to death without so much as attempting to get the group to do anything. On Fricker's account, one should tell the bystander that one blames him for his failure to attempt to unilaterally cajole, and explain why it was morally required of him to do so. This dialogue would, ideally, lead him to recalibrate his moral beliefs and come to see that he flouted an important moral requirement.

I believe that this understanding of individuals' obligations in group contexts works for most cases. That is, it works in cases where coordination of the group is at least *possible*. Where this model breaks down, though, is in cases of the kind I described in chapter 1: cases where each bystander knows for certain that attempting by themselves to ensure that the group coordinates in the right way (unilateral cajoling in Held's Subway) will be ineffectual. In that kind of case, it is not yet clear why any of the bystanders should attempt to ensure that the

group coordinates in the right way, since they know that their attempt won't make any difference to the outcome. I will be thinking about these difficult cases in chapter 4.

## 2.4 Conclusion

The goal of this chapter was to develop an understanding of three important moral concepts which are in play in Held's Subway: responsibility, blame and obligation. I argued that we should be focused on backward-looking moral responsibility, where moral responsibility for an action is to be understood as it being appropriate to take the fact that one performed the action as a basis for a moral appraisal of one's character. Responsibility understood in this way is non-gradable, making it importantly different from blame, which comes in degrees. I endorsed a Scanlonian account of blame, on which an agent is blameworthy for an action if that action reveals something about the agent's moral attitude that would impair the relationships that the agent is able to have with others. The Scanlonian view also yielded the conclusion that in order for an agent to be blameworthy for an action, they must necessarily be morally responsible for it; but the converse does not hold. An agent's action can only be taken as revealing a morally bad attitude if it is appropriate to use the fact that the agent performed the action as a basis for a moral appraisal of the agent's character. Yet it being appropriate to use the fact that the agent performed the action as a basis for a moral appraisal of the agent's character does not imply that the action reveals a morally bad attitude on the part of the agent.

In order to know whether one's action reveals a morally deficient attitude, one must first understand what the agent's obligations were in the context in question. I suggested that there are certain situations where one has moral obligations simply in virtue of the fact that

one, or a group one belongs to, as a certain capability: in other words, 'can' sometimes implies 'ought'.

This understanding of obligation allows us to see why, when there is a possibility that an agent can ensure that a group coordinates in the right way, the agent has an obligation to attempt to do so. What the discussion so far doesn't account for is why any of the bystanders in Held's Subway should have unilaterally cajoled even in cases where they knew for sure that doing so would be ineffectual.

Those questions will be asked later in the thesis. Before turning to them, I will consider a different candidate answer to the Responsibility Question in Held's Subway. This answer comes from Virginia Held herself, and will therefore be called 'Held's Answer'.

### **Chapter 3: Held's Answer**

The purpose of this chapter will be to examine Virginia Held's answer to the Responsibility Question in Held's Subway. Held suggests that the bystanders bear *collective responsibility* for the victim's death.<sup>42</sup> Thus, the process of evaluating Held's suggestion will require a detailed enquiry into what could be meant by 'collective responsibility'. I will ultimately argue that, although the kind of collective responsibility which Held endorses is appropriate in some settings, it is not applicable in Held's Subway.

#### **3.1 Random Collections**

After wondering about how responsibility should be assigned in Held's Subway, Held writes, 'I think that in such a case we would hold the random collection morally responsible for its failure to act as a group'.<sup>43</sup> To understand this, we must first understand what Held means by 'the random collection'.

Explanation of this term comes earlier in the paper, where Held tells us that 'I shall mean by a "random collection of individuals" a set of persons distinguishable by some characteristics from the set of all persons, but lacking a decision method for taking action that is distinguishable from such decision methods, if there are any, as are possessed by all persons'.<sup>44</sup> The group of bystanders on Held's Subway therefore constitutes a random collection because the bystanders share the characteristic of being on the same subway at the same time, but they lack any procedure for deciding how to act. Google is not a random

---

<sup>42</sup> Such a view, or something very similar to it, has also been endorsed by, among others, May (1992), p.116-117, Copp (1991), p.75-76 and Isaacs (2011), p.147-154.

<sup>43</sup> Held (1970), p.477

<sup>44</sup> Ibid. p.471

collection, because, although the people at Google share a common characteristic (namely, working for Google), they *do* have a distinguishable decision-making procedure – decisions are made by the board, I imagine, according to some accepted mechanism. The set containing me, the Queen of Sweden and John Travolta is not a random collection, because the three of us share no characteristic which distinguishes us from the rest of humanity (other than the fact that we make up the set I just referred to, but that fact is too trivial to make any difference).

There remains, though, a wide variety of groups which, on Held's account, can properly be called random collections. White people, people who like heavy metal music, South Americans, and East Londoners are all presumably random collections, since they all share a distinguishing characteristic yet lack a joint decision-making procedure. As Stanley Bates points out, Held largely restricts herself to 'groups of people brought together in space and time, facing situations which seem morally to call for action by the group'.<sup>45</sup> Held's Subway deals with just such a random collection. For my purposes, I will for the moment follow Held in thinking only about these kinds of random collections.

In concluding that the random collection in Held's Subway can be held morally responsible for the death of the victim, Held thereby endorses some kind of *collective responsibility*. Therefore, to establish whether she is right I need to get clearer on the concept of collective responsibility.

---

<sup>45</sup> Bates (1971), p.344

### 3.2 Some Important Distinctions

Perhaps even more so than 'responsibility' and its cognates, the term 'collective responsibility' has been put to a vast array of different uses by philosophers and non-philosophers alike. Caution is therefore required before reaching any conclusions about whether Held's suggestion is viable.

As I see it, there are three main ideas which have been referred to as 'collective responsibility' by philosophers. We must make sure to keep these concepts separate in our minds, and I will go through them now, using nomenclature which I hope is helpful in this regard.

First, there is *shared responsibility*. Shared responsibility is what results when a series of people independently contribute to some result, and none of them are aware of how their action contributes to the wider result. In other words, it is the responsibility arising from sets of uncoordinated actions. Each person whose action contributed to the wider result shares responsibility for that result.

Larry May has the following elucidating example of shared responsibility.

*Consider a situation in which a number of people know an important fact about a person being investigated by the government, and each decides to say nothing to the investigators. If each person on his or her own decides to keep quiet, and if each person's silence is necessary for the deceit to occur, then it appears that each person is equally to blame for whatever harm the deceit causes. Since each contributes to the silence and each person's contribution is necessary, there is no*

*reason, so this argument would suggest, to treat one person differently from another. This seems especially plausible since none of them is aware of what the others are doing or not doing.*<sup>46</sup>

Aggregate action on the part of agents, where their action is a necessary part of some wider effect, results in shared responsibility on the part of each agent, when the agents aren't to know that their action is contributing to this wider effect.<sup>47</sup> Note that in the case of shared responsibility, it is the *individuals* who are bearers of moral responsibility.

It's important to note the requirement that the individual actors must not have any knowledge of the fact that their action is contributing to some wider effect, and, we can add, that they couldn't have been expected to know this. This means that shared responsibility doesn't arise when the actors either knew or *should have known* that their act was part of a wider set of acts which was contributing to some effect. This is especially important for many of the problems facing humanity today, for example climate change. If I decide to drive my gas-guzzling 4x4 around the corner to the shop, rather than walking, this may seem like an independent, uncoordinated action on my part. I might then plead that I at most bear shared responsibility for the effects of climate change, perhaps thinking that this is a less serious accusation than others which might be levelled at me.

However, such a plea on my part would be unwarranted. This is because it's common knowledge that individual actions which pollute the environment are part of a wider trend

---

<sup>46</sup> May (1992), p.115.

<sup>47</sup> All uses of the word 'action' here could be replaced by 'inaction' and the point would still stand.

which, cumulatively, is having disastrous effects on our planet. Even if I don't know the science, I can reasonably be expected to know that my action is contributing to a wider effect. Therefore, I don't bear shared responsibility – I bear some other kind of responsibility.

A second concept which has been referred to by philosophers as 'collective responsibility' is *corporate responsibility*. 'Corporate responsibility' refers to cases where the bearer of responsibility is a group of people. The following commonplace utterances all rely on corporate responsibility being possible: 'Facebook is to blame for the rise in online bullying'; 'the Cabinet is responsible for the war in Iraq'; 'the Navy deserves praise for its handling of the crisis'. Those who endorse the possibility of corporate responsibility therefore commit themselves ontologically to the existence of a group, over and above its members, which is capable of being responsible for actions. Proponents of corporate responsibility have several difficult questions to answer.

Firstly, they must give an account of what a group is, and how groups can act. They must also give an account of which kinds of groups can bear moral responsibility or, if they think that all kinds of groups can bear moral responsibility, they must justify that position.

Secondly, they must explain how groups can be rational agents, since most of our intuitions about responsibility presuppose that a necessary condition of an agent's being responsible is that she is rational and was able to act according to her intentions. If the proponent of corporate responsibility wants to deny this assumption, they must provide a good argument against it.

Thirdly, they must provide an account of how, if at all, corporate responsibility is distributed to the individuals of the group. Steven Sverdlik writes that ‘with corporate responsibility the group is treated as being distinct from its members and responsibility for wrongdoing is attributed to it. If one supposes that corporate responsibility is possible then it is an open question whether the individuals in the group are also responsible for the outcome or whether the group as such is alone responsible’.<sup>48</sup>

These questions, though difficult, are by no means insurmountable, and some philosophers, notably Peter French and Philip Pettit, have provided convincing answers to them in endorsing corporate responsibility.<sup>49</sup>

The philosophical debate around the idea of collective responsibility has been significantly hampered by the unfortunate practice of using ‘collective responsibility’ to refer to corporate responsibility, at the same time as others have used it to refer to something else. Several notable thinkers have written articles whose title is ‘Collective Responsibility’, but whose topic is corporate responsibility – this has muddied the waters.<sup>50</sup>

The final concept often referred to using the term ‘collective responsibility’ is, I think, the one most worthy of the name, and I will therefore call it ‘collective responsibility’. The main idea behind collective responsibility is that individuals can be responsible for the acts or omissions of groups of which they are members. Note the crucial feature – it is individuals, not groups, who bear collective responsibility. Sverdlik writes, ‘collective responsibility is the idea that

---

<sup>48</sup> Sverdlik (1987), p.62.

<sup>49</sup> See French (1979), French (1982) and Pettit (2007).

<sup>50</sup> See, for example, Cooper (1968) and the response by Downie (1969). Also Narveson (2002).

individual persons within a group are responsible for an outcome produced collectively. That is, responsibility is apportioned to individuals and to them alone'.<sup>51</sup> The important (and controversial) aspect of collective responsibility is that individual members can be responsible, not only for the acts or omissions that they themselves do, but the acts or omissions of the group as a whole. This feature of collective responsibility is what has led people to condemn it as attempting to hold people responsible for other people's actions – something which many people take to be unacceptable to any system of morality. Thus, in his famous article, H.D. Lewis calls the notion of collective responsibility 'barbarous', arguing that 'no one can be responsible, in the properly ethical sense, for the conduct of another...belief in "individual" as against any form of "collective" responsibility is quite fundamental to our ordinary ethical attitudes'.<sup>52</sup> It's worth dwelling on Lewis' paper, because his aversion to what he understands as collective responsibility can help us to become clear on what exactly collective responsibility is, and, just as importantly, what it is not.

Lewis appears to have two main complaints against collective responsibility. The first is borne of his absolute certainty that only individuals can be the bearers of moral responsibility. In addition to the sentence I gave above, he writes right at the start of the paper 'responsibility belongs essentially to the individual'.<sup>53</sup> This, however, is a point upon which Lewis and the proponent of collective responsibility agree; as I have explained, collective responsibility ascribes responsibility only to individuals. This, then, turns out not to be a point of contention.

---

<sup>51</sup> Sverdlik (1987), p.62

<sup>52</sup> Lewis (1948), p.3

<sup>53</sup> Ibid. p.3

Lewis' second complaint appears to be that he takes collective responsibility to entail that, in his words, 'we are all involved in the sins of all'.<sup>54</sup> He is concerned that, on collective responsibility, each person bears responsibility for all sins that happen everywhere in the world and that 'in respect to properly moral worth, there is nothing to choose between the lives of individuals'.<sup>55</sup> The thought seems to be that, if collective responsibility is permitted, we lose all sense of being individual moral agents, and individual good and bad becomes overrun by a global good and bad, in which we are all implicated.

This is a caricature. It is easy to see, given Lewis' exaggerated understanding of collective responsibility, why he thought it was barbarous – if he were right about what collective responsibility meant, it would indeed be barbarous. The idea that I bear responsibility for things done on the other side of the planet from me, of which I had no knowledge, is absurd. Fortunately, this is not what collective responsibility says. Collective responsibility advocates the much milder position that I can be responsible for outcomes brought about by groups of which I am a member. What is required to be a member of a group in the appropriate way is, of course, in need of explanation, but it is not the sense of membership on which I am a member of the group containing all people in the world.

The idea of collective responsibility as individuals being responsible for outcomes brought about by groups they belong to is the conception put forward by Christopher Kutz. In a discussion of the bombing of Dresden during World War 2, Kutz cites soldiers involved in the bombing who reported afterwards that they felt a great sense of responsibility for the

---

<sup>54</sup> Ibid. p.5

<sup>55</sup> Ibid, p .5

annihilation of the city, even though their individual contributions paled in comparison to the total damage done. Kutz proposes that the responsibility they feel is a result of what he calls the Complicity Principle, which he says is well-grounded in our intuitions, ethical practices, and psychologies.

***The Complicity Principle:** (Basis) I am accountable for what others do when I intentionally participate in the wrong they do or the harm they cause. (Object) I am accountable for the harm or wrong we do together, independently of the actual difference I make.<sup>56</sup>*

Where Kutz uses the word ‘accountable’, I propose that we substitute ‘morally responsible’, understood as I have described so far. I will discuss Kutz’s position in greater detail in the next chapter. For now, it will suffice to note that collective responsibility allows that we can be individually morally responsible for outcomes that are brought about by groups to which we belong.<sup>57</sup>

Shared responsibility, corporate responsibility, collective responsibility: three distinct concepts which have, at various points by various people, been referred to using the term ‘collective responsibility’. They must be kept separate in our minds. The following table summarises the findings of this section.

---

<sup>56</sup> Kutz (2000), p.122. Note that Kutz distinguishes between the ‘basis’ and the ‘object’ of accountability, the basis being the reason I am to be held accountable, and the object being what it is I am accountable for.

<sup>57</sup> Kutz’s position is what Julia Nefsky has called ‘Strong Participation’. See Nefsky (2015), p.250-253.

|  | <b>Shared responsibility</b> | <b>Corporate responsibility</b> | <b>Collective responsibility</b> |
|--|------------------------------|---------------------------------|----------------------------------|
| Is the individual the primary bearer of responsibility?  | Yes                          | No, the <i>group</i> is.        | Yes                              |
| Does the individual responsibility arise from an action (or inaction) that the individuals did <i>together</i> ? | No                           | N/A                             | Yes                              |

This discussion was borne out of a desire to evaluate Virginia Held's position that we can hold the random collection in Held's Subway responsible for the death of the victim. I concluded that Held endorsed some kind of collective responsibility. I must now ask which of the three concepts outlined in this section Held had in mind.

### 3.3 Which Collective Responsibility?

Which of the three species of responsibility did Held mean to ascribe to the random collection? Sverdlik (who is well aware of the distinctions) appears to think that Held is talking

about collective responsibility – he calls Held’s paper ‘one of the most discussed articles defending the notion of collective responsibility’.<sup>58</sup>

However, I think that it is apparent in Held’s paper that she is not talking about collective responsibility as Sverdlik and I are understanding the term – rather, she is talking about *corporate* responsibility. Here is a selection of sentences from the paper providing evidence for this view.

*The question I wish to examine...is whether a random collection of individual persons can ever count as a collectivity or group capable of action for the performance or non-performance of which the group may be held morally responsible.*<sup>59</sup>

*Even though we may be willing to allow that collectivities can act, the questions whether they can also be held responsible and whether they can be subject to moral judgement, remain. I assume an affirmative answer to both.*<sup>60</sup>

*It seems that we can sometimes conclude that the judgement “Random collection of individuals R is responsible for not doing A” is valid.*<sup>61</sup>

---

<sup>58</sup> Sverdlik (1987), p.63, footnote 4.

<sup>59</sup> Held (1970), p.471

<sup>60</sup> Ibid. p.472

<sup>61</sup> Ibid. p.478

It's apparent from these remarks, with their emphasis on the *group's* or the *collective's* responsibility, that Held has in mind corporate responsibility. That is, when she says that it's possible for the random collection to be morally responsible for the non-performance of an act, she means that the group *itself* can be held morally responsible. It's only a secondary question for Held whether or not this responsibility distributes to members of the group as well.<sup>62</sup>

Now, using the term 'corporate responsibility' to evaluate Held's Answer may be misleading, because Held's random collections are not incorporated at all – that is one of their key features. For clarity, then, I will adopt a different nomenclature: I will from now on refer to corporate responsibility as 'group responsibility', to emphasise that Held's claim is that the group is responsible *qua* group, even though the group is not a corporation.

Now that I have a better understanding of the claim Held is making, I can begin to evaluate it. The question has become whether or not there can be *group moral responsibility* on the part of a random collection when the random collection fails to perform some act. Held thinks that this is possible. I will argue that it isn't.

Before giving the argument, I should first acknowledge that I haven't shown that Held understands the term 'moral responsibility' in the same way which I understand it following my enquiry in chapter 2. Held may have an entirely different conception in mind. However, Held's idea that there is group moral responsibility in the case of Held's Subway is nevertheless

---

<sup>62</sup> Ibid. p.480 for discussion of this. Held thinks that the random collection's moral responsibility is distributive – that is, responsibility is distributed to every member of the random collection, though perhaps 'in significantly different proportions'.

a putative answer to my question of who bears moral responsibility in Held's Subway. I may not be asking precisely the same question which Held is asking, but her answer is still a candidate answer to my own question, and therefore deserves attention here.

### 3.4 Against Held's Answer

As I have mentioned, some of the most plausible defences of group responsibility have been produced by Peter French and Philip Pettit. In their work, they focus on showing how certain groups should be properly treated as moral agents capable of acting according to intentions, and should therefore be included in the group of entities which can bear moral responsibility. Pettit argues that groups can be counted as autonomous agents when they have the ability to form and re-form 'action-suited' desires and beliefs, and then act in order to satisfy those desires based on those beliefs.<sup>63</sup> French argues that groups can become moral agents when they have both 'an organizational or responsibility flow chart that delineates stations and levels within the corporate power structure' and 'corporate decision recognition rules'.<sup>64</sup> Those sympathetic to Held's view, then, might appeal to the work done by French and Pettit as evidence that group moral responsibility is possible.

However, there is a major difference between the groups which French and Pettit are chiefly concerned with and those which Held is concerned with. French and Pettit are thinking about complex, highly organised groups which have established decision procedures, whereas Held is thinking about 'random' groups of people who have never met. Held's groups are thrust together, with no hierarchy or decision-making process, into a situation which seems to call

---

<sup>63</sup> Pettit (2007), p.178

<sup>64</sup> French (1979), p.212

morally for action. The question is whether this difference between the kinds of groups which Pettit and French discuss, and those which Held discusses, is relevant to the question of group responsibility.

Pettit certainly thinks so. He writes of 'unincorporated groups' that 'the members [can't] be held collectively, as distinct from individually, responsible, that is, responsible as a collection or group. *For the collection that they constitute, being unincorporated, is not an agent*'.<sup>65</sup> For Pettit, being 'unincorporated' precludes a group from being able to bear moral responsibility for things. I agree with Pettit, and will now offer an argument for this position.

1. Recall that my understanding of moral responsibility, established in chapter 2, is that for someone to be morally responsible for an action or inaction is to be such that it is appropriate to take the performance or non-performance of that action as a basis of a moral appraisal of that person's character. Keeping that in mind, I offer the following assumption.

**Assumption 1:** only moral agents can be subject to moral appraisal.

As support for this assumption, consider that we don't take the actions of beings which we consider to be non-moral-agents as bases for appraisals of their character. To return to the child doodling in crayon on the wallpaper; we don't take that act to be in any way indicative of a bad moral character on the part of the child. The same goes for profoundly impaired humans, and inanimate objects such as chairs. The chair's collapsing when I sit on it is not used

---

<sup>65</sup> Pettit (2007), p.195, my emphasis.

to morally appraise the chair (although it may be taken as evidence of the abilities of the manufacturer).

Rejecting this assumption would mean that there are some things in the world which are not moral agents but *are* subject to moral appraisal for their actions. The contradictory nature of that position is, I hope, sufficient evidence that assumption 1 is correct.

2. **Assumption 2:** An important feature of being a moral agent is the ability to have intentions.

Being a moral agent requires the ability to have intentions. Note that the converse does not hold – one can have intentions yet fail to be a moral agent. The toddler who draws on the wall in crayon presumably intends to do so, but this doesn't make her a moral agent.

This understanding of agency is based upon Davidson's account.<sup>66</sup> Again, it is lent plausibility by the implausibility of its negation. That is, to assert that something can be a moral agent even though it is unable to have intentions seems highly implausible.

3. **Assumption 3:** For groups of people to have a collective intention, certain features must be present, similar to those outlined by Bratman and Tuomela and Miller.

---

<sup>66</sup> Davidson (2001), p.45-46

Much work has been done in determining whether groups can have intentions, and how this could be so. Some of the leading accounts are given by Michael Bratman, Raimo Tuomela and Kaarlo Miller. I will briefly summarise their positions here. The important thing to note is that, according to these accounts, groups can have intentions, but this is no trivial matter. It is much more complicated than, for example, every member of the group having the same intention individually.

On Bratman's account, in order for a group of two people (you and I) to have a shared intention to J, a series of conditions must be met.<sup>67</sup> Firstly, it must be the case that I intend that we J and that you intend that we J. Secondly, it must be the case that I intend that we J in accordance with and because of your intention that we J, and that you intend that we J in accordance with and because of my intention that we J. This condition guards against my forcing you to J under duress (even though, unbeknownst to me, you actually intend to J), since my forcing you under duress to J is not our jointly intending to J. Thirdly, I must intend that we J according to 'meshing subplans'<sup>68</sup> of our separate intentions to J. That is, we must intend that our separate intentions to J will align on certain important aspects of the plan. For instance, if J is painting our house, we must intend that our separate plans will mesh on, for instance, the colour we are going to paint it, and who is going to paint which section (but we needn't intend that the plans mesh, for instance, on which shop I am going to buy the paintbrushes from). Finally, all of the above must be common knowledge to both of us. Note

---

<sup>67</sup> Bratman (1993), p.106

<sup>68</sup> Ibid. p.106

also that Bratman is only talking about cases involving two people – one can imagine how complicated the account would be in the case of larger groups.<sup>69</sup>

Tuomela and Miller argue that for a group to have a joint intention to X, it must be the case that each of the individual members of the group have the appropriate ‘we-intention’, which is an intention to the effect that ‘we will X’ or ‘the group will X’.<sup>70</sup> In order for a member of the group to have a we-intention, they must satisfy three conditions. First, she must intend to do her part in X. Second, she must believe that a sufficient number of members of the group will do their part in X (that is, she must believe that X will be brought about). Finally, she must believe that the belief that X will come about (that is, that a sufficient number of members of the group will do their part) is held by all members of the group.<sup>71</sup>

In making assumption 3, I am suggesting that, in order for groups to have intentions, something like the accounts put forward by Bratman and Tuomela and Miller must be satisfied. That is, it is not enough simply for members of the group to all have the same intention; a more stringent set of conditions must be satisfied.

4. The final assumption of the argument is as follows.

**Assumption 4:** Held’s random collections are unable to satisfy the conditions for group intention outlined in assumption 3.

---

<sup>69</sup> Bratman explicitly rules out the notion that shared intention can be achieved simply by two people having the same intention as one another. See Bratman (1993), p.103.

<sup>70</sup> Tuomela and Miller (1998), p.370

<sup>71</sup> Ibid. p.375

It may be possible for certain groups to satisfy the requirements for group intention. These groups would then genuinely have intentions, which is different from saying that all of their members independently have the same intention. However, from the way Held describes her random collections, I argue that they aren't capable of satisfying the requirements.

Recall that Held describes random collections as 'lacking a decision method for taking action that is distinguishable from such decision methods, if there are any, as are possessed by all persons'. That is, random collections have no system by which to arrive at group decisions, or to decide which actions the group needs to take. They are made up of people thrust together by circumstance with no prior agreement on any 'meshing subplans' or 'we-intentions' that the others might have. They can also have no common knowledge of the kind which is required by both accounts given above. They couldn't possibly know what the other members of the random collection know or believe, since they have never met them before.

5. If assumptions 2,3 and 4 are correct, it follows that Held's random collections are not moral agents. Working backwards: if assumption 4 is correct, random collections can't meet the requirements for group intention set out by Bratman, Tuomela and Miller; if in addition assumption 3 is correct, random collections can't have intentions; if in addition assumption 2 is correct, random collections are not moral agents.

6. If assumption 1 is correct then, from step 5, random collections can't be subject to moral appraisal, since only moral agents can be subject to moral appraisal.

7. From 6, random collections can't be morally responsible, because for an agent to be morally responsible requires that the agent can be subject to moral appraisal.

If this argument works, it shows that Held was wrong to say that the random collection can be held morally responsible for failing to help the victim in Held's Subway. She was wrong, because to be morally responsible for an action (or inaction) means that one's performance (or non-performance) of the action can be taken as a basis for a moral appraisal of one's character, but the argument has shown that random collections can't be subject to moral appraisal and are therefore not the kind of thing that can be morally responsible for anything.

Kutz disagrees with my assumption 3. That is, he questions the idea that for a group to have a collective intention they must satisfy some demands of the kind stipulated by Bratman, Tuomela and Miller. Kutz says that no such complicated set of requirements must be met – rather, groups having intentions is a much simpler matter than all that.

As evidence for this position, Kutz offers the following example. Suppose that you and I are having a picnic and it begins to rain. Without stopping to hold a debate on what we should do about this, I grab the basket of food and run to the car. Seeing me do this, you grab the blanket and follow me. Kutz says that we have the joint intention of 'saving the picnic' in this situation. He writes that 'if we do both act with participatory intentions, then we will have jointly intentionally saved the picnic though neither had formed an intention to save the picnic in the light of expectations about the other's intentions'.<sup>72</sup>In other words, in order for us to be

---

<sup>72</sup> Kutz (2000), p.92

properly described as having a joint intention, it's not the case that we need to have knowledge of what the other person believes, and beliefs about whether they are going to do their part. Instead, we simply need to act with participatory intentions, and believe that it's *not impossible* that others will join us and do their part too.

If Kutz were right, then it would be damaging for my argument because it would show that random collections needn't satisfy requirements like Bratman's, Tuomela's and Miller's in order to have intentions. They may be able to have intentions in some other way, in which case they might properly be regarded as moral agents in certain situations, and appropriate subjects of moral appraisal.

However, I think it is questionable whether Kutz actually has identified an alternative way for groups to have intentions. It seems to me that what Kutz has described, in the 'saving the picnic' case, is a situation in which two individuals have the same intention as one another: to save the picnic with the help of another. This is very different from saying that the group they make up has an intention. Two MPs from opposing parties may both have the intention to reduce poverty, and they may both hope that the other will join them in trying to achieve this. Supposing that they both in fact do work to try and reduce poverty, this isn't enough to say that the group made up of those two politicians has the intention to reduce poverty. More is needed than this: something that binds the two of them together and allows them to form a joint plan – something, in short, like the proposals of Bratman, Tuomela and Miller.

I conclude that Held's answer to the question of who is responsible for the death of the victim in Held's Subway is not right. However, there is another approach to the problem which says

that the random collection is not responsible for not helping the victim; rather, it is responsible for *failing to come together as a group which is capable of joint action*. Held makes this point when she says '[a] random collection can be held morally responsible for failing to make a decision on which action to take – for failing, that is, to adopt a decision method'.<sup>73</sup> In other words, the random collection isn't morally responsible for not attempting to save the victim; rather, it's responsible for not forming itself into an organised group which could have done something to help.

Now, the argument just given rules this out, for the same reason: that random collections aren't the kind of entities that can be morally responsible for things. Nevertheless, this comment from Held does seem intuitively appealing. There is a great desire to say that the people on Held's Subway should have come together and formed a group capable of action. Given that the random collection isn't morally responsible for this failure, who is?

The answer, I believe, is that the *individuals* are morally responsible. Each of the five bystanders in Held's Subway bears individual moral responsibility for their failure to become a group capable of collective action. This point is well made by Tracy Isaacs, who writes of the failure to form a group, 'we need to see this as a failure of individuals because it takes place in the absence of a moral agent'.<sup>74</sup>

---

<sup>73</sup> Held (1970), p.479

<sup>74</sup> Isaacs (2011), p.147

### 3.5 Conclusion

The goal of this chapter was to evaluate a candidate answer to the question of who is morally responsible for the death of the victim in Held's Subway. The suggestion was that there is group moral responsibility: the random collection *qua* random collection is morally responsible. I rejected this view because a random collection is not the kind of thing that can be morally responsible.

It is *individuals*, not the group, who bear moral responsibility in Held's Subway, and they are morally responsible for failing to attempt to bring the group together. Each of the individuals should have attempted to bring the group together; talk of the group itself being responsible is a distraction.

In the next chapter I will turn to the important objection which has been lurking beneath the surface of the thesis so far. That is, what are we to say about Held's Subway if cajoling on the part of any particular bystander would have been ineffectual and everyone knows it? How can we hold the individuals responsible for not cajoling if, had they done so, it would not have made a difference to the outcome?

## **Chapter 4: What If I Couldn't Have Made a Difference?**

If unilateral cajoling by any particular bystander in Held's Subway would have resulted in the bystanders coming together as a group capable of preventing the assault, then each bystander is to blame for failing to do so. As I argued in chapter 2, this derives from the obligation that each member of a group has to attempt to ensure that the group coordinates in the right way when the group has a chance to do something good or prevent something bad. The bystanders are blameworthy for their inaction, because their inaction reveals an attitude of theirs towards others which would impair their relationships with other people; namely, they aren't sufficiently concerned with the welfare of others. They flouted a moral obligation they had to cajole.

What are we to say, though, of cases where unilateral cajoling would have been unsuccessful? If each bystander knew for sure that unilateral cajoling would not have made any difference, can we still apportion blame to them in the same way? Or would they be blameless in this case for doing nothing?

In the last chapter I surveyed how Virginia Held thought responsibility should be ascribed in this kind of case – Held thinks that the random collection *itself* is morally responsible. If I am right to disagree with Held, it follows that the moral responsibility, and so the blame, if there is any in cases like this, must be borne by individuals. But how can the individuals be to blame if their individual actions wouldn't have made a difference?

Although it might seem far-fetched in Held's Subway that any of the bystanders could have known for certain that unilateral cajoling would have failed, in other real-world cases it is far

more plausible. For instance, to anticipate chapter 5 slightly, if I am a member of a random collection capable of successfully campaigning for traffic control measures in my village, it is quite possible that I might know for sure that unilateral cajoling on my part will have no effect. Furthermore, it is ultimately these real-life cases which I hope this thesis will shed some light on, so it is important to consider a version of Held's Subway in which everyone knows that unilateral cajoling won't work.

As I see it, there are three versions of Held's Subway which possess this feature. I will go through them one by one, and will be asking which, if any, of the bystanders is blameworthy in each. I will ultimately defend the intuition that all of the bystanders are blameworthy in every case, but to widely varying degrees, and widely varying reasons. That is, I will be defending the intuition that failure to cajole on the part of the bystanders in Held's Subway is blameworthy, even if they knew that cajoling by themselves wouldn't have made a difference.

#### 4.1 Case 1: When Others Might Also Cajole

The first case is one in which, although unilateral cajoling would not be successful, nevertheless there is a chance that others will also simultaneously cajole. We should imagine for this scenario that each of the bystanders must make a quick decision, upon seeing the assault begin, whether or not they are going to stand up and attempt to rouse the group, thereby demonstrating their willingness to form part of a rescuing group. That is, the other bystanders are not cruelly indifferent to the victim's fate – it is uncertain whether they will attempt to rouse the group. If two or more bystanders cajole, the group is formed, since there will be at least two willing rescuers, and it was specified in chapter 1 that two or more rescuers

could overcome the attacker. Can we say that any of the bystanders are to blame for not cajoling in this instance?

We might think that we can hold the bystanders blameworthy for not cajoling in this case on purely consequentialist grounds. This is the line that Shelly Kagan has taken.<sup>75</sup> In what Kagan calls *threshold cases*, most acts of a given kind make no difference to an outcome, but after a certain amount of such acts have been performed, the *next* one pushes over a threshold, triggering some bad event (or a good event, in the case I am considering). Kagan's example is of people buying chickens from a supermarket. My buying the first chicken of the day doesn't contribute to the cruelty involved in raising and slaughtering chickens for the supermarket – the delivery of chickens has already been made, and 25 chickens arrived this morning which were raised in awful conditions and slaughtered. That can't be changed, and my buying a chicken can't turn back time. However, once the 25<sup>th</sup> chicken of the day is sold, let's imagine, the supermarket must put in a new order with the chicken farm that 25 more chickens are required. This order triggers the slaughtering of 25 more chickens, which is a terrible thing.<sup>76</sup> In this case, the person who buys the 25<sup>th</sup> chicken *does* make a difference: consequentialism can easily condemn her act, since it would have been a better state of affairs had she not bought the chicken. She would have forgone chicken for dinner, but there wouldn't have been another slaughter.

In fact, Kagan continues, the act of buying a chicken only makes a difference if the total number of chickens bought from the supermarket that day is a *multiple* of 25. For only then is

---

<sup>75</sup> Kagan (2011)

<sup>76</sup> Of course, the example is simplified and the numbers no doubt wildly inaccurate, but the point stands.

it the case that, for each of the people who bought a chicken, the outcome would have been different if they had done otherwise. If 50 people buy a chicken then, had any one of them *not* bought a chicken, only one new order would have been made, whereas in the event two new orders were made. Suppose that 53 people bought a chicken; now each purchaser can truthfully say that, had he not bought a chicken, the outcome would have been the same, since two new orders would still have been made that day, given how everyone else chose to behave.

The step Kagan then takes is to appeal to *expected outcomes*. If it is the case that my act will only make a difference if the total number of chickens is a multiple of 25, then I have a  $1/25$  chance of my act of buying a chicken making a difference. I have a  $24/25$  chance of my act not making any difference – the total number of chickens bought could be  $25n+1$ ,  $25n+2$  etc., up to  $25n+24$ . I must conduct a calculation of the expected outcome of my act of buying a chicken. In this case, I have  $1/25$  chance of my act bringing about a terrible event (the slaughtering of 25 more chickens) and a  $24/25$  chance of my act not making a difference. The small chance of the bad event happening, Kagan stipulates, is enough for consequentialism to tell you not to do it. The expected utility of your act of buying a chicken is negative.

We might take a Kagan-style approach to argue that one should cajole in this version of Held's Subway where each bystander is unsure whether their individual act of cajoling would have made a difference. That is, we could appeal to the expected utility of one's action. The approach will be the reverse, so to speak, of Kagan's chicken case. Now, there will be a small chance that my act will bring about something good, and a large chance that my act will make no difference.

I stipulated that the bystanders would not respond if only one person cajoled, but that if at least two people cajoled, a group would form that was capable of stopping the attack.<sup>77</sup> From the point of view of a bystander trying to decide what to do, then, the following table applies.

| Total number of cajolers | Will my cajoling make a difference? |
|--------------------------|-------------------------------------|
| 1                        | No                                  |
| 2                        | Yes                                 |
| 3                        | No                                  |
| 4                        | No                                  |
| 5                        | No                                  |

If only he cajoles, nothing will happen, so his cajoling won't make a difference. If 3,4, or all 5 of the bystanders cajole, his cajoling will be superfluous: the outcome would have occurred whether or not he cajoled. Only in the case where exactly two people cajole (him and one other), can we say that his act made a difference; without him, the outcome would have been different.<sup>78</sup>

---

<sup>77</sup> At the very least, the capable group will comprise the two cajolers, since, by cajoling, they have demonstrated their willingness to join in. I leave open the possibility that the two cajolers are so convincing that some of the remaining three bystanders are moved to throw their weight behind the group. The point is that one person cajoling alone will be insufficient.

<sup>78</sup> Recall that, for this version of the case, the bystanders must all make their decision on whether or not to cajole simultaneously.

Following Kagan's line, the bystander should then calculate the expected utility of his action. Let  $X$  be the probability that he will be one of exactly two cajolers. His calculation will therefore be:

$X \times$  (*bring about a group capable of saving the victim's life*)

+

$(1-X) \times$  (*minor inconvenience of cajoling without effect*)

Given the extremely large upside of organising a group capable of saving the victim's life, it is likely that the expected utility calculation will tell him to stand up and cajole in this case even for small values of  $X$ .

This is a Kagan-style argument that consequentialism can explain why a failure to cajole on the part of the bystanders is blameworthy in this version of Held's Subway. Even though their unilateral cajoling wouldn't have made a difference, there is a chance that they would have been one of a duo of cajolers who roused the group to action. The potential upside is sufficiently large that, despite the small chance of it occurring, it is enough to outweigh the minor convenience of remaining in one's seat. Therefore, each bystander should have cajoled, and they are blameworthy for not having done so.

I find Kagan's argument persuasive, but its weakness is that it doesn't generalise well when the numbers change in the example. As Julia Nefsky has written of the chicken example:

*Kagan is right that the numbers as he has them entail a negative expected utility for purchasing a single chicken. However, the numbers clearly need not be as he imagines them to be. And it is only because he has made the numbers match up—for some  $N$ ,  $N$  purchases causes  $N$  additional chickens to be hatched and raised who wouldn't otherwise have been—that he gets the results he wants: automatic negative expected utility. There is no reason to think that the numbers must match up in this way.<sup>79</sup>*

The same is true of this version of Held's Subway. Although Kagan's reasoning works well in the case as I have described it, it may give less intuitively appealing results in slightly different situations. Consider the following variation.

Suppose instead that there were not five bystanders on a subway, but 100 bystanders on a commercial aeroplane. Suppose also that it wasn't a person being beaten and strangled to death, but, say, a 30-year-old man who had fallen and become trapped under a refreshment trolley. It will take five people to lift the trolley off his leg, and if it doesn't get lifted off within 10 minutes then he will suffer permanent damage and have to walk with a stick for the rest of his life. Finally, suppose that all of the passengers are somewhat apathetic – not actively malevolent, but fairly lazy and not particularly concerned about the man's leg. Imagine that if one person stands up and tries to cajole the other passengers into helping him lift the trolley, nothing would happen. However, if two people cajole, a capable group will come together.

---

<sup>79</sup> Nefsky (2012), p.369-370, her emphasis.

By the same reasoning as before, a passenger considering whether or not she should cajole faces the following expected utility calculation, with  $X_1$  the probability that she will be one of exactly two people to stand up.

$X_1 \times$  (*bring about a group capable of saving the victim's leg*)

+

$(1-X_1) \times$  (*minor inconvenience of cajoling without effect*)

Two things strike me about this case. First, passengers who don't attempt to rouse a capable group in this situation are, it seems to me, blameworthy. While it is not a life at stake, there is clearly the potential to avoid a serious harm befalling the unfortunate man if a capable group is formed. Not to cajole in this instance reveals a certain morally deficient attitude, and may make one reconsider the kind of relationships one could have with such a person.

Secondly, it isn't obvious that Kagan's expected utility calculus approach gives an answer in accordance with this intuition.  $X_1$  is likely to be extremely small, and almost certainly smaller than  $X$ . When there are 100 people, the chances of exactly two of them attempting to rouse the group will be lower than if there are only five people. Also, the upside is not as large as in Held's Subway; a life is not about to be lost. This makes it at least questionable whether the upside can overcome, as it were, the very low chance of one's action making a difference in

bringing it about. People may disagree, but I suspect that the expected utility calculation in this instance will not unequivocally say that one should cajole.

Consequentialists may respond to this in two ways. First, they may bite the bullet and say that in fact, and contrary to the intuition I have put forward here, there is no obligation to cajole in the aeroplane case. Second, they may maintain that the potential upside of saving the man from a lifetime of walking with a stick actually *is* enough to overcome the very low chance of making it happen. In other words, the expected utility calculation will in fact give the 'right' answer here.

I find the first response unconvincing. To suggest that passengers who do nothing to attempt to get a group together to save the man are blameless is, I maintain, to set a very low bar for morality. The second response can be argued about. The point I want to make is that it is at least unclear whether the consequentialist can get the 'right' answer in this case. Indeed, even if consequentialism does get the 'right' answer here, different variations in the numbers will throw up cases which are even more problematic for Kagan's expected utility approach.

My question is how we can blame a bystander who fails to cajole in the version of Held's Subway where it requires two people to cajole in order for a capable group to form. Kagan's expected utility approach gives a plausible answer to this question, but his reasoning doesn't generalise well to other cases. Therefore, I must look for a better explanation of the intuition that the bystanders must cajole – one that applies to a wider range of situations where one's action almost certainly won't make a difference to the outcome.

Nefsky offers a convincing account. Her crucial step is to reject the assumption that helping to bring about an outcome requires making a difference. Such an assumption is often made in everyday life: for instance, we say that there's no point in voting because my vote isn't going to make a difference to the outcome. However, Nefsky points out that there is value in *helping*; making a difference isn't the only thing that has moral worth. Her definition of what it means to help bring about an outcome is as follows.

*Suppose your act of X-ing could be part of what causes outcome Y. In this case, your act of X-ing is non-superfluous and so could help to bring about Y if and only if, at the time at which you X, it is possible that Y will fail to come about due, at least in part, to a lack of X-ing.<sup>80</sup>*

Let us apply Nefsky's model to the version of Held's Subway currently under consideration. X-ing, in this case, is cajoling the other bystanders into forming a group capable of stopping the attack. Y is a capable group forming. Nefsky's idea, then, is that cajoling is non-superfluous and so could help to bring about a capable group forming if and only if, at the time at which the bystander cajoles, it is possible that no capable group will form due to a lack of cajoling.

I suggest that, if cajoling would help to bring about a capable group in this way, then each bystander has an obligation to cajole, so long as the upside is sufficiently large, and the cost of cajoling is sufficiently small for each bystander. In Held's Subway, these two conditions are likely to be met, since the upside is very large (forming a group capable of saving the victim's

---

<sup>80</sup> Nefsky (2016), p.2753

life), and it doesn't take much to cajole the other bystanders. Note that there is no requirement to cajole if two others have already done so. Cajoling after two others have already done so would be superfluous because, at the time at which the third person cajoles, it is not possible that a capable group will fail to form, since two people cajoling would be sufficient to bring the group together. This seems right – it would be strange to think that somebody has an obligation to cajole even after the good outcome has already come about. Again, the crucial aspect of Nefsky's view is that at no point does her model mention making a difference. Whether or not your action will make a difference is irrelevant to whether or not it will help, and is therefore not the main factor contributing to whether or not you have an obligation to do it.

Nefsky's idea also deals well with the aeroplane case. Here, cajoling one's fellow passengers to form a group capable of lifting a trolley off the man's leg will help to bring this outcome about so long as there is a chance that the outcome will fail to come about due to a lack of cajoling. I suggest that you have an obligation to help to bring the good outcome about, and so therefore an obligation to cajole, unless a sufficient number of people have cajoled already (and supposing cajoling doesn't impose a substantial cost on you, and that the upside is sufficiently large).<sup>81</sup>

As Nefsky mentions, her account also provides good reasons why people should vote in elections.<sup>82</sup> There is only a miniscule chance that general elections will turn on a single person's vote. For that to happen, the result must either be a tie or a one-vote-victory. This

---

<sup>81</sup> Note that while I am talking about obligations, Nefsky only talks about having *reasons* to help. See Nefsky (2016), p.2744-2745 for discussion.

<sup>82</sup> Nefsky (2016), p.2754

fact has led people to suggest that there is no point in voting, since one's vote is not going to make a difference. Nefsky's account provides a response to this line of argument. Nefsky is willing to admit that one's vote is not going to make a difference, but she nevertheless says that one has a reason to vote (I would say one has an obligation to vote) if there is a chance that a bad candidate will win as a result of not enough people voting for a good candidate. The important feature of the case is that the outcome is yet to be determined, and it may go the wrong way if not enough people do a certain action. The focus on whether or not one's vote will actually turn the election is a distraction.

Nefsky acknowledges that 'it is hard to shake the idea that helping requires making a difference'.<sup>83</sup> However, she explains that this is because in typical, *non*-collective impact cases, it is *true* that helping requires making a difference. That is, if you can't make a difference in these ordinary cases, you can't help. In these ordinary cases, at least one of three preconditions for helping doesn't hold: either your act will not be part of what causes the outcome; or your act will be superfluous to the outcome; or there will be no chance of the outcome failing to come about. Nefsky gives the example of a student taking a final exam for a pass/fail course. She realises that, due to her other coursework, even if she gets a zero on the final exam, she will still pass the course. She reasons that taking the exam will make no difference to whether or not she passes, and that therefore taking the exam will not help her to pass the course. The third condition of helping isn't satisfied: there is no chance of the outcome (her passing) failing to come about.

---

<sup>83</sup> Ibid. p.2758

Our intuitions about helping and making a difference are forged on examples of this kind: typical, *non*-collective impact cases. Nefsky alerts us to the fact that the intuition does not transfer over correctly into collective impact cases.

I conclude that in this version of Held's Subway, where unilaterally cajoling won't make a difference, but one might be part of a duo of cajolers who are successful, the bystanders are blameworthy for not cajoling. This is best explained by acknowledging Nefsky's discovery that the ability to make a difference is not the only consideration that can confer obligations on people. The ability to *help* can also mean that one has an obligation to help.

#### 4.2 Case 2: Five Bad People

The case just discussed left open the possibility that some of the other bystanders would also cajole. In cases 2 and 3, this possibility is removed. The question in cases 2 and 3 is why a bystander might be blameworthy for failing to cajole in a version of Held's Subway in which she knows for sure that none of the other bystanders are going to help – that is, when there is no chance of a capable group forming. I will defend the intuition that all of the bystanders are blameworthy for failing to cajole in such cases.

The first of these kind of cases is one in which all five of the bystanders are mean, selfish people who simply don't care about their fellow passenger, and everyone knows it. None of them are even contemplating getting up to cajole, but even if one of them did, it would be useless because the rest of them aren't going to help anyway.

How should blame be apportioned in this case? I think, persisting with the Scanlonian view of blame which I considered in chapter 2, that the best way to think about this is to say that each of the individuals is to blame for failing to help the victim because their non-act shows something about their attitudes towards others that impairs the relations that others can have with them. Certainly, if I was to subsequently meet any of the five bystanders, I would not wish to enter in to any kind of relationship with any of them, knowing that they had such callous attitudes towards their fellow passenger. Their non-action is blameworthy because of what Scanlon calls the ‘meaning’ of their non-action: ‘the significance, to the agent and others, of the agent’s willingness to perform that action for the reasons he or she does’.<sup>84</sup>

Now, we may not be able to tell, as observers, what the meaning of the non-action for each of the Five Bad People was. Because cajoling wouldn’t have made a difference, we might not be able to distinguish them from good people who just know that they are surrounded by bad people who will not help – such a case is the subject of the next section. Nevertheless, the grounds for their blameworthiness is the meaning of their action. Lamentably, these grounds may be inaccessible to us, with the result that we are unable to actually blame them ourselves.

### 4.3 Case 3: One Good Person, Four Bad People

#### *4.3.1 Why the One Good Person might be blameworthy*

The final version of the case, which I will call One Good Person, Four Bad People, involves one bystander who is caring and eager to help the victim of the assault, and four others who are uncaring and selfish, just like all the bystanders were in case 2. Imagine that the Good Person

---

<sup>84</sup> Scanlon (2008), p.4

is trying to decide whether to cajole the others into forming a group capable of saving the victim. She knows that her fellow bystanders are all Bad People, and that they aren't interested in helping, and won't take part in any rescue mission no matter what the Good Person does. Who, if anyone, is blameworthy in this version of Held's Subway?

The first thing which strikes me is that the Four Bad People are all blameworthy for the same reason as the Five Bad People were all blameworthy in case 2. That is, they are blameworthy because their non-act revealed an attitude of theirs which impairs the relations which others can have with them; namely, their callous indifference to the victim's plight.

So much for the Four Bad People. The interesting feature of case 3 is whether or not the One Good Person is blameworthy too. I think that most people will intuitively say that she is *not* blameworthy. It's not her fault, some might argue, that she is on a subway with Four Bad People. Not only does she know for sure that her cajoling won't make a difference, but she also knows that it won't even help, because there is no chance that a capable group is going to form. Cajoling in this case seems like a waste of time, and we can't hold people to be blameworthy for failing to do something which is a waste of time.

Despite this, I am going to defend the intuition that the One Good Person is in fact blameworthy for not cajoling in this version of the case, although she bears blame to a much smaller degree than the other bystanders.

To motivate this view, consider the following example.

### Sharks vs. Jets

Suppose that A is taking a walk along a river with some friends who all happen to be part of a notorious gang called the Sharks (A isn't part of the gang, but she occasionally likes to spend time with these particular Sharks). As they make their way along the river bank, they come across a drowning man. As they get closer, they are able to clearly identify the man as B. B is known to both A and her friends as one of the ring leaders of a rival gang, the Jets, with whom the Sharks have a long-running and hateful feud. It's widely known that B cannot swim, and it is obvious that he will drown in a few minutes without assistance. It is also obvious that no one person out of the group on the bank could haul B out (he is a big man), but that two or more of them, each grabbing a part of B's clothing, could successfully get him out. Now suppose that A knows for sure that, because of the bitterness of the relations between Sharks and Jets, and the uncommonly vehement hatred of Jets which these particular Sharks have, nothing she says or does will persuade any of her Shark friends to help her haul B out of the river. She knows that the Sharks would all be happy to let B drown. To remove any element of personal danger for A, assume that her Shark friends are very fond of her and, if she did attempt to persuade them to help rescue B, they would not become aggressive or angry, but would simply politely refuse.

Suppose, knowing this, that A decides to keep her head down and walk along in silence, while her Shark friends gleefully celebrate the impending death of B. My question is: is A blameworthy for that? I have a strong intuition that she *is*. There is something morally repugnant, I suggest, about A's behaviour. On the Scanlonian view, her non-act reveals an

attitude of hers (which particular attitude it is is not yet clear), I maintain, that impairs the relations which others can have with her.

I think that Shark vs. Jets reveals the intuition that I am seeking to defend more effectively than One Good Person, Four Bad People, and therefore I will work with Sharks vs. Jets for the remainder of this section.<sup>85</sup> I want to defend the intuition that A is blameworthy for her non-actions. The problem is that it is extremely hard to explain *why* A is blameworthy. Several moral theories and thinkers who have written about this topic are unable to explain this intuition, as I will now show.

#### 4.3.2 Some ethical theories which can't explain the intuition

##### *Nefsky*

In section 4.1 I argued that Nefsky's idea of the importance of *helping* in collective impact cases, rather than *making a difference*, can explain why bystanders in case 1 ought to cajole, and can be blameworthy if they do not. This is because by cajoling they can help to bring about a good outcome which may yet not come about through a lack of cajoling.

However, Nefsky's idea isn't going to work in Sharks vs. Jets. That is because in Sharks vs. Jets there's no possibility of the good outcome coming about. As Nefsky writes, 'in collective

---

<sup>85</sup> The reason, I think, that the Sharks vs. Jets case makes things clearer than the One Good Person, Four Bad People case is that, despite having all the same features, those features are more believable in Sharks vs. Jets than they are in One Good Person, Four Bad People. For instance, it is much easier to imagine that A would know for sure that the Sharks wouldn't help than it is to imagine that the One Good Person would know for sure that the Four Bad People wouldn't help, because A knows the Sharks and their history with the Jets. It is also easier to imagine that A could safely attempt to persuade the Sharks without risk of injury to herself; in One Good Person, Four Bad People, our perception may be blurred by the thought that, if the One Good Person cajoles, the attacker may turn on her (though as I stipulated in chapter 1, for the sake of this example, we should assume that this would not happen).

impact cases, there being other people who will or might act in the relevant way is necessary for being able to satisfy the conditions for helping'.<sup>86</sup> In *Sharks vs. Jets*, there are no such other people.

### *Consequentialism*

Act consequentialism can't handle these cases in the way we would like it to. According to act consequentialism, an act is right if it is the act out of all the available acts which brings about the best state of affairs. In *Sharks vs. Jets*, A's attempting to persuade the Sharks to help save B would not have made a difference to the resulting state of affairs – B would have died anyway, since none of the Sharks would have helped. If A's cajoling would not have made a difference to the resulting state of the world, how can consequentialism prescribe that she should have cajoled? In case 1, one possible solution was Kagan's appeal to expected utilities, and the importance of the *chance* that one's action will make a difference. However, in *Sharks vs. Jets*, that isn't going to work, since there is *zero* chance that A's cajoling would have made a difference.

Another thinker who has sought to provide a version of consequentialism which can accommodate collective impact problems is Felix Pinkert. Pinkert finds Kagan's approach unconvincing, because it only applies to cases where agents are unsure whether or not their action will make a difference – it doesn't adequately deal with cases where agents know that they cannot bring about a different outcome by acting differently. Pinkert's example is of two factories operated by polluting owners who are such that if either one of them stopped

---

<sup>86</sup> Nefsky (2016), p.2755

polluting and started producing cleanly, the clean producer would go out of business, and the environment would still be badly damaged: in short, unilaterally defecting to produce cleanly would bring about a worse state of affairs than continuing to pollute.<sup>87</sup> Pinkert also stipulates that both polluters are uncooperative – were polluter A to start producing cleanly, polluter B would not cooperate, and would continue to pollute, and vice versa. The case presents a problem for consequentialism, since by consequentialism's standards, both polluters act rightly, since if either of them acted differently, a worse state of affairs would be brought about.

Pinkert suggests the following solution to the problem.

***Modally Robust Act Consequentialism:** An agent ought to act optimally in the actual world, and be such that for all possible combinations of the actions of other agents, if that combination were instantiated, she would act optimally in these circumstances.<sup>88</sup>*

The idea is that an agent can be blameworthy, even if they act rightly according to consequentialism in the actual world, if they are such that they would act sub-optimally in other possible situations. Pinkert's thought is that this allows us to blame each of the factory owners, because even though they both acted rightly given what the other owner actually did, they both *would have* been uncooperative had the other one chosen to produce cleanly. They

---

<sup>87</sup> Pinkert (2015), p.973

<sup>88</sup> Ibid. p.982

would have caused a sub-optimal result in a slightly different scenario in which the other factory owner acted differently.

However, Pinkert's idea isn't going to be able to explain my intuition that A ought to have cajoled in Sharks vs. Jets. Pinkert will say that A did *right* by consequentialism in the actual world by not cajoling: cajoling would have made no difference to the resulting state of the world, so A was correct to save herself the trouble and remain silent. However, it's also true that *had* one or more of the Sharks been receptive to the idea of saving B, A would also have acted optimally in *that* scenario. If one of the Sharks had, amazingly, decided to cooperate, A would also have cooperated, since she is a good person and wanted to save B. Thus A satisfies Pinkert's Modally Robust Act Consequentialism. Pinkert is unable to find her blameworthy for not unilaterally cajoling.<sup>89</sup>

I conclude that there has not yet been any rendering of consequentialism which is able to explain the intuition that A ought to unilaterally cajole in Sharks vs. Jets.

### *Parfit*

In 'Reasons and Persons' Derek Parfit sought to give an account of why acts may be wrong even though they were right by consequentialist standards. He wrote that

*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*

---

<sup>89</sup> Pinkert may be untroubled by this: he may not share my intuition that the One Good Person is in some sense blameworthy for not unilaterally cajoling.

*be what someone ought to do, because it is one of a set of acts that together benefit other people.*<sup>90</sup>

We might take this to be pointing us in the direction of why it was wrong for A to fail to cajole in Sharks vs. Jets; her act might have been right in and of itself (after all, her acting differently would not have made any difference to the outcome), but yet it may have been wrong because it was part of a set of acts which harm someone – namely, the acts (or non-acts) of the group on the riverbank which resulted in nobody intervening to help B. Perhaps this allows us to say that A is blameworthy.

However, I don't think that Parfit's idea can give us this much. It is true that A's act was part of a set of acts which together harm the victim. But this would be true of *any act which A had performed*. If A had cajoled, her act would still have been part of a set of acts which harmed B, and so would still, by Parfit's analysis, have been wrong. A is simply in the unfortunate position that, whatever she does, her act will be considered wrong on Parfit's view. This doesn't give her a reason to cajole, any more than it gives her a reason to do anything else.

In a later, unpublished manuscript, Parfit actually rolls back from his 'Reasons and Persons' position on sets of acts.<sup>91</sup> His later position is to endorse subjective act consequentialism, on which the 'subjectively right act is whatever will have greatest *expected* goodness'.<sup>92</sup> It's apparent that this later position also can't explain why A should have cajoled. There is no expected goodness to be brought about by A's cajoling – it would have had no effect on the

---

<sup>90</sup> Parfit (1984), p.70

<sup>91</sup> Parfit (1988), p.1

<sup>92</sup> Ibid. p.4, his emphasis

outcome, and A knows this. By remaining silent A maximised expected utility, since she saved the wasted energy of attempting to persuade the Sharks.

I conclude that none of the above ethical theories can explain the intuition that A should have attempted to persuade the Sharks to help her save B. I nevertheless reaffirm my intuition that she *should* have cajoled, and that she is blameworthy for remaining silent and walking on by.

Now, the preponderance of theories just reviewed which give the result that A is *not* to blame for failing to attempt to persuade the Sharks might suggest that my intuition is bogus. Perhaps I am wrong to feel that A is blameworthy for her non-action in Sharks vs. Jets – perhaps I am morally confused in some way about the case. Nevertheless, I find myself unable to shake the intuition that A is blameworthy for her non-action, and will proceed with trying to defend it. What follows will therefore fall short of an unassailable argument that will convince those who doubt that my intuition is in fact correct. Rather, I offer the coming pages merely as a defence which those who already agree with my intuition can use to tentatively attempt to convince others.

#### *4.3.3 An ethical theory which might explain the intuition*

I think that the intuition that A should have cajoled the Sharks might be explained by Christopher Kutz's theory of complicity, which I introduced briefly in chapter 3. Kutz's view is that we can be accountable for the actions that groups of which we are part perform if we intentionally participate in those actions, even if our contribution to the group action is negligible. We are what Kutz calls the *inclusive authors* of acts which groups of which we are part perform, and can therefore be held accountable for those group actions.

Kutz gives the example of two people having a picnic together. One of them is getting the drinks and food out of the car, and the other one goes on ahead and picks somewhere to put down the blanket. Unbeknownst to the first person, the second person puts down the blanket on a bed of flowers, destroying them. Kutz says that, although the first person hasn't done anything wrong in this case, he is accountable for the destruction of the flowers as an inclusive author of that act. He willfully took part in the picnic, and so will be in some way accountable when the gardener comes over to demand an explanation.<sup>93</sup>

Might Kutz's notion of inclusive authorship explain why A acted wrongly by failing to cajole? We might be able to say that she is accountable for the act done by the group of which she is a part. She is part of a group which performed the non-action of failing to come to B's aid, and thus she is in some way accountable for that.

The problem with this line of argument is that A did not act with participatory intention. She did not willfully join the others in the collective failure to help – rather, she did it *unintentionally*, wishing that she were not doing it. On Kutz's view, she would therefore not be appropriately called an inclusive author of the act. 'Inclusive accountability', says Kutz, 'is based on the teleological...relations between group members' intentions and the collective act. We are properly held accountable for the actions of groups...in which we participate, because these actions represent our own conception of our agency and our projects'.<sup>94</sup> Since

---

<sup>93</sup> See Kutz (2000), p.143.

<sup>94</sup> Ibid. p.140-141

the group's non-act did not reflect A's intentions, she can't be seen as an inclusive author of the non-act. Therefore, she can't be held accountable or blameworthy for the non-act.

However, I think Kutz's theory nevertheless gets us closer to what is objectionable about A's failure to even attempt to persuade the Sharks to help. She may not be properly seen as an inclusive author of the non-act, but she nevertheless did not *sufficiently distance herself from the non-act*. Her failure to make any kind of protestation suggests a level of toleration of the group's failure which is blameworthy. Kutz writes that the conception of inclusive authorship 'expresses what we desire, what we will *tolerate*, and what we *believe*'.<sup>95</sup> I propose that what is objectionable about A's failure to cajole is that it suggests a certain toleration on her part of the group's indifference to B's plight.

Of course, any sort of protestation on A's part would, as I have specified, have had no effect. What I am suggesting, then, is that what was required of A was the performance of a *symbolic gesture* – an act which would have had no effect other than to distance her from the group's failure. That such a symbolic gesture can be required of a group member is hinted at by Kutz in the following passage.

*The nascent sense of common venture I am invoking can be linked to a second basis of accountability for unstructured collective harms: symbolic, or character-based accountability...In overdetermined contexts, agents can have reason to refrain from participating in a harm, not because of the relation between this choice and*

---

<sup>95</sup> Ibid. p.141, my emphasis.

*an actual outcome, but because of what the choice symbolizes in their characters and commitments... agents who distinguish themselves from other participants demonstrate a commitment to the value of the lives of those they harm...The motive in these cases is not...causal. That is, agents need not believe that unilateral nonparticipation will lead others to follow. Rather, they choose to act as a way of expressing meaning.*<sup>96</sup>

The notion of disassociating oneself from collective acts is endorsed by Howard McGary.<sup>97</sup> McGary gives the following two conditions under which 'moral agent X can be held morally liable for faulty practice P'.

*(1) X knows or should have known about P.*

*(2) X identifies or has solidarity with those who engage in P or X does not sufficiently disassociate himself from P or X's failure to disassociate from P was not a part of a reasonable strategy to prevent further or greater harm.*<sup>98</sup>

On McGary's view, then, even if one has no solidarity with those who engage in P it is still required that one sufficiently disassociate oneself from the action if one wants to avoid blame. The moral agent must do something to distance himself from the morally bad act (or

---

<sup>96</sup> Ibid. p.190, his emphasis.

<sup>97</sup> The notion is also given an excellent and thoroughgoing examination by Thomas Hill Jr. (1979). It is endorsed as well, though given a less thorough investigation, by John Lucas (1993), p.85: 'A man may reckon that the chances of his being heard and heeded are nearly nil, but he still owes it to himself to witness the truth and to stand up for things he believes in, even though ineffectually'. Such a view is also hinted at by Peter French (1974), p.283.

<sup>98</sup> McGary (1986), p.162, my emphasis.

non-act). McGary acknowledges that what is required to sufficiently disassociate oneself will vary from case to case. In some cases, simply speaking up will suffice, but in others more emphatic actions will need to be taken, such as direct intervention, and a refusal to accept any benefits which come your way as a result of the bad practice P.

My suggestion is that A is blameworthy for not attempting to persuade the Sharks to help because by failing to speak up she does not sufficiently distance herself from the non-action of the group. Note that it might well be the case that performing a symbolic gesture of the kind I am suggesting will not entirely absolve A of blame.<sup>99</sup> She may still be blameworthy, but to a lesser extent than if she had not performed the symbolic gesture.

It might be objected that to disassociate yourself in this way, even when doing so can have no effect on the outcome, is to adopt what Kutz has called 'an annoying priggishness or high-mindedness'.<sup>100</sup> It might appear that someone who seeks to disassociate herself is overly concerned by her own moral purity, and places too much importance on maintaining a clean moral report card over the actual implications of her actions. This objection has strengths and weaknesses.

Firstly, the criticism is slightly misplaced in Sharks vs. Jets, since there is nothing A can do to affect the outcome, so it's not as if she must choose to keep her moral record clean over taking some practical steps. A symbolic gesture was all that was available to her. Secondly, I

---

<sup>99</sup> Juha Räikkä has argued that it is possible to distance oneself from a harmful group action, yet nevertheless still be blameworthy. Depending on what he calls the 'rule of criticism' in a society, it might be that in order to adequately oppose a practice, one must also partake in it. See Räikkä (1997).

<sup>100</sup> Kutz (2000), p.190

don't agree that a desire to avoid the blame of our fellow humans is priggish or high-minded. Being blamed for something, especially on the Scanlonian view, is no insignificant thing. It implicates who one is as a person and the values and attitudes that one has. As Scanlon himself writes, 'the realization that you have done something that gives others good reason to revise their understanding of their relationship with you, and that they also take this view, is a serious matter'.<sup>101</sup> To want to avoid being blamed is not to be excessively concerned with one's integrity, but to want to preserve the meaningful relationships that one has with others.

Nevertheless, this objection does aim at an important point, which is that it does not seem especially laudable of A if she performs her symbolic protest *solely* in order to avoid blame.<sup>102</sup> Avoidance of blame, as I have just mentioned, can be an important motivating factor for those who wish to maintain meaningful friendships with others. But a person who went through life with the ambition to avoid blame at all costs would appear to be ignoring other important reasons for action. A morally better approach to life would be to be motivated by wanting to *do the right thing*, according to one's conception of the right. In A's case, we might say that she should have placed more importance than she did on standing up to injustice, for example.<sup>103</sup>

There are similarities here with Bernard Williams' famous example of George the chemist.<sup>104</sup> George, the new PhD who is offered a job working in a chemical weapons factory, must decide whether to take the job in order to support his young family, or turn it down in the knowledge

---

<sup>101</sup> Scanlon (2008), p.157

<sup>102</sup> Thanks to Edgar Phillips for pressing this point.

<sup>103</sup> This, I think, is the point Hill Jr. makes in Hill Jr. (1979), p.97.

<sup>104</sup> Williams (1973), p.97. See also Hollis (1995) and Williams (1995) for illuminating discussion.

that a zealot will be appointed in his place who will pursue the work with a greater deal of fervor, with probably worse effects for humankind. If George's primary reason for turning down the job was that he wished to avoid the blame that would come his way in the event that the company pioneered new and deadly chemical weapons, we would perhaps not find him particularly morally admirable. Williams' suggestion is that George's *integrity* requires of him that he not take the job; given who he is – his beliefs, projects, desires etc. – he is unable to take the job.

Similarly, in *Sharks vs. Jets*, we might be concerned that A was not sufficiently moved to protest on B's behalf. On the Scanlonian view, A is blameworthy for failing to attempt to persuade the Sharks because her non-action revealed that she was not sufficiently troubled by the injustice to speak out. Such an attitude would impair the relations that one would be able to have with her. If she also acted out of a desire to avoid blame, that is understandable; but it would be less morally commendable than if she had acted from integrity.

There remains an important problem. If, as I am suggesting, the correct way to understand why A is blameworthy in *Sharks vs. Jets* is by appeal to something like Kutz's theory of complicity, it must be the case that A is *part* of the group which performs the non-action of failing to save B. If she were not part of the group, then she couldn't be complicit, and so wouldn't have to perform an action which distances her from the group's act.

In what way, though, is she part of the group? As I have mentioned, she doesn't share any participatory intention with the other group members – she doesn't *willfully* take part in the failure to save B. She also couldn't have affected the outcome – no matter what she had done

B still would have drowned. A seems to be no more part of the group than a distant explorer up a mountain, watching the scene on a television. If A is part of the group, why not the explorer?<sup>105</sup>

I suggest that what makes A, and not the explorer, a member of the group is the following counterfactual condition: if the other members of the group had had different attitudes, A might have been able to help bring about a good outcome. First, note that this is not true of the explorer. Had the Sharks changed their ways and decided that they were in fact open to the idea of saving B, the explorer still could not have had any effect on what the group did. A, however, could very well have had an effect – faced with a more cooperative group, her attempts to persuade may have been successful.

This counterfactual condition produces the result that we are all part of infinitely many groups which could bring about certain outcomes. For instance, when I am in the doctor's waiting room, amongst the magazines and several other strangers, I am part of a group which could, if the members were so inclined, form a *tableau vivant* of The Last Supper. The fact that neither I, nor any of the other people in the waiting room, have any desire to form the *tableau vivant* is, on my account, not pertinent to my being a member of the group, since on my account what is important is the counterfactual that *if* the other group members had been so inclined, *then* I could have helped bring about the desired outcome. And this is true in the doctor's waiting room; if the others had a desire to form the *tableau vivant*, I could have completed the scene by taking my place in it.

---

<sup>105</sup> Thanks to Joe Horton and Han van Wietmarschen for pointing this out.

Perhaps some people may be perturbed by this. But I do not see anything particularly disquieting about the fact that I am a member of many groups capable of performing certain actions, if the other members were so inclined. I am not suggesting that, in the doctor's waiting room, my fellow patients and I have any *obligation* to form the *tableau vivant*. If we don't do it, we have not behaved morally badly. I merely suggest that if we *could* do it, then I am a member of a group which could do it.

Of course, my view also commits me to the view that we are often members of groups which could do more morally relevant acts. These capabilities may confer certain obligations on me, although it is not clear when exactly these obligations will arise, nor how strong they will be. When I am in the pub, I am a member of a group (consisting in all the people in the pub) which could raise a great deal of money for charity if we all decided to forego our next drink and instead donate the money we would have spent to an agreed-upon good cause. Such a group act would no doubt be morally good, and there may be an obligation on me to attempt to ensure that it happens. Whether or not I have such an obligation will depend, as I have mentioned, on the personal cost for me of persuading, and the size of the good outcome that we can bring about.

#### 4.4 Conclusion

The purpose of this chapter was to ask who, if anyone, is to blame in versions of Held's Subway where each bystander knows for sure that unilateral cajoling on their part will not make any difference to the outcome. That is, if they unilaterally cajole, the victim will still die.

I have argued that, in any case like this, all five of the bystanders are to blame, although they are to blame for different reasons, and to different degrees. Those bystanders who refuse to cooperate because they are callously indifferent to the victim's plight are to blame because their non-action reveals an attitude of theirs which impairs the relationships that others can have with them. A bystander who knows that it will require at least two cajolers to make any difference might question why he is blameworthy for not cajoling – after all, his action would very probably have made no difference. I have argued that making a difference is not the morally relevant concept in a case like this; rather, one should do what one can to *help*, if there is a danger that the good outcome will not come about due to not enough people helping in the right way. Finally, even a well-meaning bystander who frustratingly finds herself surrounded by uncaring, selfish people is nevertheless blameworthy if she fails to unilaterally cajole the others into action. This is because, if she wants to avoid blame, she must sufficiently disassociate herself from the actions of the group. This intuition of mine was hopefully made more vivid through the Sharks vs. Jets example.

Of course, the degree of blame which each kind of bystander will bear is very different. I will not put forward an argument for how the types of blame should be ranked in terms of severity: I will merely state my opinion that, pre-theoretically, it seems to me that the most severe kind of blame of those which I have identified in this chapter is the type of blame one is due when one's act reveals a callous indifference towards the suffering of a fellow human, such as was revealed in case 2. The least severe kind of blame is the blame one bears for failure to disassociate oneself from a collective harm which one could do nothing to prevent – this is the type of blame appropriate for A in Sharks vs. Jets. In between these cases is the blame one bears for failing to *help* in group situations – failing to make it more likely that

some good outcome will come about, or that some bad outcome will be avoided. Others' intuitions may disagree on this ranking. The point is that, on any version of Held's Subway, if nobody attempts to cajole their fellow bystanders into action, all of the bystanders are blameworthy, for one reason or another.

## **Chapter 5: Why This Matters**

Many of the challenges we face in the modern world can only be solved with collective action. This is true of large-scale, global issues such as climate change, as well as more local issues which affect only a small community. Because of this, we need to develop a sophisticated understanding of what our individual obligations are in group situations, and what we must do in order to adequately fulfill these obligations.

Helpfully, another feature of modern life is that we are all more interconnected than we have ever been. Theoretically at least, this should make it easier for groups to form which are capable of solving the collective action problems which we face.

So far, this thesis has been concerned with the somewhat contrived example of Held's Subway, and the ethical lessons we can learn from it. In this final chapter, I will move away from Held's Subway to a more realistic scenario. I will attempt to apply the lessons learned from Held's Subway to determine what individuals' obligations are in this scenario, and who, if anyone, is to blame if a good outcome is not brought about.

### **5.1 The Busy Road**

Suppose that there is a small village through which runs a main road. Cars use this road as a shortcut to drive between two big cities either side of the village. As a result, the road is very busy, especially at rush hours, and cars drive too fast since there are no traffic-calming measures in place. There are several local amenities situated along this road, including a school.

Suppose that this situation of cars driving recklessly through the village persists for some time – say, six months. It becomes apparent that one of these days there is likely to be a car accident involving a child getting hit by a speeding car as she walks to or from school. Now let's imagine that such a tragedy does occur, and a child is hit and killed one morning.

We might ask, as we asked in Held's Subway, who is responsible and who is to blame. Again, an obvious answer is available – the person driving the car is to blame for being reckless and driving too fast. Might there also be blame on the part of anyone else?

My suggestion is that there is a random collection made up of the villagers who were aware of the danger of the speeding cars. That random collection, let us suppose, had the capability to successfully lobby their local authority to install traffic-calming measures – speeds bumps, speed cameras etc. – which would have prevented the death of the child. I will stipulate that no one villager, lobbying alone, could have successfully got these measures installed, but that had a sufficient number of them lobbied as a group, they would have got the attention of the local authority, and the measures would have been installed. Which, if any, of the villagers are blameworthy in such a case?

Firstly, let me verify that the villagers can indeed be seen as a random collection as Held has defined random collections. Recall that, for Held, a random collection is 'a set of persons distinguishable by some characteristics from the set of all persons, but lacking a decision method for taking action that is distinguishable from such decision methods, if there are any, as are possessed by all persons'. The villagers, therefore, qualify as a random collection: they

are distinguishable from the set of all persons by their shared characteristic of living in the same village, but (we can imagine), they have no particular decision method for taking action.

The villagers may be tempted to think that, since none of them acting alone could have successfully had the traffic-calming measures installed, the failure is a group failure, and it is the group which is morally responsible and blameworthy. They may take comfort in that, thinking that it is better in some way to be part of a blameworthy group than to be blameworthy oneself. However, as I argued in chapter 3, random collections are not the kind of entities which can be morally responsible or blameworthy. If there is blame in this case, it can only reside with the individuals involved.

Given that no one of the villagers could alone have successfully lobbied the local authority, the only available actions each of them had was to try to *persuade* the other villagers to join together to form a group capable of successful lobbying – the parallel of cajoling in Held's Subway. In Held's Subway, cajoling was straightforward – all the bystanders had to do was talk to the four other people who were directly in front of them, and ask them to join in. In The Busy Road, things are not so simple. The villagers are dispersed throughout the village – it's impossible to talk to them all at once. However, cajoling is still possible. They could put up flyers, for instance, advertising a meeting in the village hall to discuss the problem of the road. They could knock on doors to try to drum up support for action. The internet has made communication between disparate people extremely easy: they could make a village website, or a Facebook group and invite all the villagers to join.

If unilateral cajoling on the part of any of the villagers would have been successful in forming a group capable of action, then, straightforwardly, this is what each villager ought to have done, so long as unilateral cajoling didn't impose an excessive cost on them personally. This obligation is derived, as I argued in chapter 2, from our obligation, when members of a group capable of bringing about a good outcome (or preventing a bad one), to attempt to ensure that the group acts in the way required to bring about that outcome. Since none of the bystanders unilaterally cajoled in this case, all are to some extent morally responsible and to blame for the death of the child. To put it another way: in order to avoid blame, they needed to unilaterally cajole.

What about cases where each villager can protest, truly, that their unilateral cajoling wouldn't have made a difference to the outcome? That is, what if they knew that, for one reason or another, a group capable of getting traffic-calming measures introduced would not form as a result of their unilateral cajoling?

As I argued in chapter 4, in cases like this, all villagers are still in some way blameworthy. It may be that their non-action revealed in them a deficient moral character. They were revealed as the kind of people who cared so little about the safety of the village schoolchildren that they don't want to get involved no matter what. It may be that they placed too much value on *making a difference* – being the single person who tips the outcome from not happening to happening – instead of realising that it is also important to *help* in collective action cases. Finally, perhaps a single good villager knew for sure that the other villagers were all selfish, and wouldn't help if they cajoled. I have argued that such a villager

would nevertheless be blameworthy for failing to make an appropriate symbolic gesture to distance herself from the failure of the group.

The point of *The Busy Road* is to illustrate how the findings from Held's *Subway* translate into situations which we are more likely to come across in our everyday lives. The main conclusion is that we, as individuals, can be blameworthy for failing to attempt to bring people together into groups which have the potential to do something good, or to prevent something bad. I deliberately go further than saying that we have a *reason* to do so; rather, I suggest that we have an *obligation* to do so, provided cajoling doesn't impose an excessive cost on us personally.

I anticipate that one objection to this view will say that it entails that our moral obligations proliferate unacceptably, to the extent that we couldn't possibly fulfil them all. After all, as I mentioned, there are so many problems and challenges in the world that can only be solved by collective action. Are we to be held in some way blameworthy for all of them if we fail to attempt to rouse a group together?

My response to this is mixed. On one hand, I do think that our obligations in collective harm cases extend further than we generally take them to. On the other, I think it stops short of becoming unmanageable.

For a start, many of the world's problems have already inspired people to establish bodies dedicated to raising awareness of them and encouraging people to join the mission of defeating them. This is true of, for instance, climate change (Greenpeace, ClimateWorks,

Alliance to Save Energy), poverty (Oxfam, Make Poverty History, UNICEF) and mass human displacement (Refugees International, International Rescue Committee, Refugee One). Here, there is no lack of group capable of overcoming the problem, as there is in Held's Subway or The Busy Road. In these kinds of cases, cajoling to bring a group together is not required. In these kinds of cases, our obligations are simpler: give what we can in order to help ensure that these organisations run as well as they need to to solve the problems.

Additionally, as I have mentioned, the obligation to cajole comes with the condition that doing so doesn't impose an excessively high personal cost on the cajoler. In this thesis I haven't gone into detail about what such a condition means in practice, but I suggest that it rules out the obligation to cajole if this will impede your ability to carry out other moral obligations you have, such as to care for your children or not to gratuitously cause somebody pain. Now, in most cases, I suggest that cajoling will not impinge on such other moral obligations. In Held's Subway, for instance, the cost of cajoling is very low, and amounts to simply expending the effort required to speak up, as well as any fear or embarrassment that comes with speaking in front of a group. In The Busy Road, the cost might be slightly higher: perhaps one would have to give up an evening or a weekend in order to knock on doors, or to spend a certain amount of time a week maintaining a website. I nevertheless maintain that the personal cost of cajoling will often be small, and so we have more obligations to speak up than perhaps we think we do. Groups of people have the potential to do great things together, and to prevent terrible things from happening. We must acknowledge the obligations that this capability puts on each of us.

## Bibliography

- Bates, S. (1971). 'The Responsibility of "Random Collections"'. *Ethics* 81 (4), pp. 343-349.
- Bratman, M. (1993). 'Shared Intention'. *Ethics* 104 (1), pp. 97-113.
- Collins, S. (2017). 'Filling Collective Duty Gaps'. *The Journal of Philosophy* 114 (11), pp. 573-591.
- Cooper, D. (1968). 'Collective Responsibility'. *Philosophy* 43 (165), pp. 258-268.
- Copp, D. (1991). 'Responsibility for Collective Inaction'. *Journal of Social Philosophy* 22 (2), pp. 71-80.
- Davidson, D. (2001). 'Essays on Actions and Events'. *Oxford University Press*.
- Dietz, A. (2016). 'What We Together Ought to Do'. *Ethics* 126 (4), pp. 955-982.
- Doan, M. (2016). 'Responsibility for Collective Inaction and the Knowledge Condition'. *Social Epistemology* 30 (5-6), pp. 532-554.
- Downie, R. (1969). 'Collective Responsibility'. *Philosophy* 44 (167), pp. 66-69.
- Feinberg, J. (1968). 'Collective Responsibility'. *The Journal of Philosophy* 65 (21), pp. 674-688.
- French, P. (1974). 'Morally Blaming Whole Populations'. In: Held, Morgenbesser and Nagel (eds.) 'Philosophy, Morality and International Affairs'. *Oxford University Press*.
- French, P. (1979). 'The Corporation as a Moral Person'. *American Philosophical Quarterly* 16 (3), pp. 207-215.
- French, P. (1982). 'Crowds and Corporations'. *American Philosophical Quarterly* 19 (3), pp. 271-277.
- Fricke, M. (2016). 'What's the Point of Blame? A Paradigm Based Explanation'. *Noûs* 50 (1), pp. 165-183.

- Gibbard, A. (1965). 'Rule-Utilitarianism: Merely an Illusory Alternative?'. *Australasian Journal of Philosophy* 43 (2), pp. 211-220.
- Hart, H.L.A. (1968). 'Punishment and Responsibility: Essays in the Philosophy of Law'. *Oxford University Press*.
- Held, V. (1970). 'Can a Random Collection of Individuals be Morally Responsible?'. *Journal of Philosophy* 67 (14), pp. 471-481.
- Hill Jr, T. (1979). 'Symbolic Protest and Calculated Silence'. *Philosophy & Public Affairs* 9 (1), pp. 83-102.
- Hollis, M. (1995). 'The Shape of a Life'. In: Altham, J.E.J and Harrison, R. (eds.) 'World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams', pp. 170-184. *Cambridge University Press*.
- Isaacs, T. (2011). 'Moral Responsibility in Collective Contexts'. *Oxford University Press*.
- Jackson, F. (1987). 'Group Morality'. In: Smart, Pettit, Sylvan & Norman (eds.) 'Metaphysics and Morality: Essays in Honour of J.J.C. Smart'. *Blackwell*.
- Jaspers, J. (2000). 'The Question of German Guilt'. E.B. Ashton (trans.). *New York: Fordham University Press*. Originally published as 'Die Schuldfrage' (1947).
- Kagan, S. (2011). 'Do I Make a Difference?'. *Philosophy & Public Affairs* 39 (2), pp. 105-141.
- Kutz, C. (2000). 'Complicity'. *Cambridge University Press*.
- Lewis, H.D. (1948). 'Collective Responsibility'. *Philosophy* 23 (84), pp. 3-18.
- Lucas, J.R. (1993). 'Responsibility'. *Oxford University Press*.
- May, L. (1992). 'Sharing Responsibility'. *The University of Chicago Press*.
- McGary, H. (1986). 'Morality and Collective Liability'. *The Journal of Value Enquiry* 20 (2), pp. 157-165.
- Narveson, J. (2002). 'Collective Responsibility'. *Journal of Ethics* 6 (2), 179-198.

- Nefsky, J. (2012). 'Consequentialism and the Problem of Collective Harm: A Reply to Kagan'. *Philosophy & Public Affairs* 39 (4), pp. 364-395.
- Nefsky, J. (2015). 'Fairness, Participation and the Real Problem of Collective Harm'. In: Timmons, M. 'Oxford Studies in Normative Ethics', volume 5, pp. 245-271. *Oxford University Press*.
- Nefsky, J. (2016). 'How you can help, without making a difference'. *Philosophical Studies* 174 (11), pp. 2743-2767.
- Parfit, D. (1984). 'Reasons and Persons'. *Oxford University Press*.
- Parfit, D. (1988). 'What We Together Do'. *Unpublished Manuscript*.
- Pettit, P. (2007). 'Responsibility Incorporated'. *Ethics* 117 (2), pp. 171-201.
- Pinkert, F. (2015). 'What If I Cannot Make a Difference (and Know It)'. *Ethics* 125 (4), pp. 971-998.
- Räikkä, J. (1997). 'On Disassociating Oneself from Collective Responsibility'. *Social Theory and Practice* 23 (1), pp. 93-108.
- Rosen, G. (2004). 'Skepticism about Moral Responsibility'. *Philosophical Perspectives* 18 (1), pp. 295-313.
- Rovane, C. (2014). 'Forward-Looking Collective Responsibility: A Metaphysical Reframing of the Issue'. *Midwest Studies in Philosophy* 38 (1), pp. 12-25.
- Rubenstein, E. (2015). 'What is it Like to be a Robot?'. *Unpublished BPhil Thesis, Oxford University*.
- Scanlon, T.M. (1998). 'What We Owe to Each Other'. *The Belknap Press of Harvard University Press*.
- Scanlon, T.M. (2008). 'Moral Dimensions'. *The Belknap Press of Harvard University Press*.
- Singer, P. (2011). 'Practical Ethics'. Third edition. *Cambridge University Press*.

- Sverdlik, S. (1987). 'Collective Responsibility'. *Philosophical Studies* 51 (1), pp. 61-76.
- Tuomela, R. and Miller, K. (1998). 'We-Intentions'. *Philosophical Studies* 53 (3), pp. 367-389.
- Williams, B. (1973). 'A Critique of Utilitarianism'. In: Smart, J.J.C and Williams, B. 'Utilitarianism: For and Against', pp. 75-150. *Cambridge University Press*.
- Williams, B. (1995). 'Replies'. In: Altham, J.E.J and Harrison, R. (eds.) 'World, Mind and Ethics: Essays on the Ethical Philosophy of Bernard Williams', pp. 185-224. *Cambridge University Press*.
- Zimmerman, M. (1985). 'Sharing Responsibility'. *American Philosophical Quarterly* 22 (2), pp. 115-122.