

**Are Individual Citizens Morally Responsible for
Policy Outcomes?**

Morten Højer Jensen

Submitted in accordance with the requirements for a degree of
Doctor of Philosophy

The University of Leeds

Faculty of Arts, Philosophy

IDEA CETL

January 2017

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

© 2017 The University of Leeds and Morten Højer Jensen

The right of Morten Højer Jensen to be identified as Author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988.

Abstract

This project concerns one overall question: are individual citizens morally responsible for policy outcomes? The aim of this project is to answer this question, in order to resolve a tension between the commonly held ideal that policy outcomes represent the will of the people on the one hand, and a seeming reluctance to hold our fellow citizens morally responsible for these outcomes on the other. In order to resolve this tension, I examine various accounts of moral responsibility to see whether the individual citizen is responsible in light of these. I focus primarily on whether the individual can be morally responsible in light of her political participation via her voting action. Firstly, I examine whether the individual is morally responsible in light of her direct contribution to the voting outcome. I conclude that she is not, because she fails to make a relevant causal contribution. I then examine indirect accounts (mainly shared responsibility), i.e. accounts of moral responsibility which do not require that the individual makes a direct contribution. I ultimately show that none of the examined accounts are successful. Therefore, I develop an account which can be successfully applied, based on observations made throughout this project. Specifically, I argue that the individual can be morally responsible for policy outcomes, if she performs her unilateral part in constituting or sustaining the particular project which brings about the policy outcome. She does this – roughly – if she through her voting action was an interdependent part of the project that brought about the outcome. She is an interdependent part even in the event where she fails to make an actual contribution to the voting outcome and thus the policy outcome itself. Lastly, I apply this account to a high stakes just war scenario, and show that it explains our intuitions of responsibility.

Contents

1	Introduction.....	7
1.1	Introduction	7
1.2	Chapter Outline	9
1.3	Unpacking the question.....	13
1.4	Why it is an interesting question	20
1.5	Summing up	24
2	Direct individualist moral responsibility.....	24
2.1	Introduction	24
2.2	Responsibility concepts	25
2.3	Substantive accounts of moral responsibility	28
2.4	Conditions for moral responsibility, and the proposed strategy ..	37
2.4.1	Attributability as control.....	38
2.5	Summing up the chapter	46
3	Causal condition.....	47
3.1	Introduction	47
3.2	Causal responsibility as a necessary condition for direct attributability responsibility, and the problem of omissions.....	48
3.3	How causal responsibility can be undermined.....	59
3.3.1	Introducing symmetrical and asymmetrical overdetermination 60	
3.3.2	Voting – a threshold case	64
3.3.3	Voting - a pre-emption case.....	66
3.4	Summing up the chapter	71
4	Cognitive condition.....	75
4.1	Introduction	75
4.2	Cognitive condition	78
4.2.1	Ignorance undermines reasons responsiveness.....	82
4.2.2	Summing up so far.....	87
4.3	Awareness of policy outcomes and culpable ignorance.....	88

4.3.1	Culpable ignorance.....	89
4.4	Cognitive conditions and voting	98
4.5	Summing up this chapter	104
5	Rounding off all the previous chapters on direct Moral Responsibility	105
5.1	Direct moral responsibility	105
5.2	Anticipating indirect accounts of moral responsibility.....	112
6	Indirect accounts of moral responsibility.....	113
6.1	Introduction	113
6.2	Simple reductionist account of collective responsibility	117
6.3	Affective identification and moral responsibility	122
6.4	Authorization, group membership, and disassociation obligations	128
6.5	Summing up this chapter	138
7	Joint action and complicity	139
7.1	Introduction	139
7.2	The complicity principle.....	142
7.3	Joint action	145
7.4	Applying and critiquing Kutz' account	152
7.4.1	Complicity and voting	153
7.4.2	Is joint action necessary for shared responsibility?	156
7.4.3	The problem with 'participatory intention'	161
7.5	Complicity and gradability	168
7.6	Summing up this chapter	172
8	Moral membership	173
8.1	Introduction	173
8.2	Observations from the previous chapters.....	178
8.3	Contribution and moral membership.....	189
8.3.1	Interdependence as a morally significant contribution	191
8.3.2	Interdependence and voting	196

8.3.3	The problems with the threshold voting case as understood thus far: 198	
8.3.4	Voting reconceptualization	200
8.4	Moral membership, the full account.....	208
8.4.1	Testing and critiquing this moral membership.....	213
8.5	Summing up this chapter	216
9	Conclusion.....	217
9.1	Introduction	217
9.2	Summing up the previous analysis	217
9.3	Moral membership and non-combatant liability in just war	219
9.4	Moral membership grounds non-combatant liability	222
9.4.1	Conclusion.....	229
10	Bibliography	230

Are Individual Citizens Morally Responsible for Policy Outcomes?

1 Introduction

1.1 Introduction

States can cause enormous amounts of harm, and states are recognized internationally as legal entities which can be held responsible for these harms, to the extent that this responsibility can ground compensatory duties. Further, it seems true that all harms caused by a state must at some point originate in the actions of certain people, and in democratic states, it is just as clear a fact that the actions of *the people* are a necessary condition for these harms to occur. This relationship is most direct, and most clear in to the voting habits of the citizens, who in a real sense, decide which candidates or policies win the day. Perhaps because of this relationship, there exists in western democracies a widespread ideal that the outcomes of the democratic process either represent, or ought to represent the will of the people. Further, from a common sense perspective, it seems to be the case that if someone's will is appropriately represented in some outcome, then the author should at least share in the praise and blame attached to this outcome, and share the associated benefits and burdens. Curiously enough then, there also exists a significant common reluctance to hold our fellow citizens accountable for said outcomes to the extent that even mild resentment towards people's seemingly immoral voting habits is uncommon. Therefore there is an apparent tension, if not inconsistency, between these observations. In the following, I will resolve this tension. Specifically, I will examine whether it is

possible to establish an account which can justify holding the individual citizen morally responsible for policy outcomes. Indeed, I will produce such an account which I will show can be successfully applied to the voting case. Specifically, I will show how citizens who voted for a given candidate (or party or policy), can be morally responsible for the ultimate policy outcomes this candidate brings about.

The paradigm example of a citizen, who is intuitively accountable for a policy outcome, is here the Nazi sympathizer who strongly supported the Nazi regime, and voted for the NSDAP in 1933 with the reasonable belief that the NSDAP winning would lead to the systematic extermination of millions of innocents. Whether any person actually voted with this foresight is for the sake of this project of less importance. However, most people would on the face of it agree that if any ordinary citizen is accountable for the harms through their political participation, it would be this person. Therefore this project will try to uncover the theoretical underpinnings of this intuition, and examine how or whether they can be generalized. While doing so, I will also examine the extent to which a successful account of moral responsibility grounds imposing certain sanctions on the individual, or otherwise leads to certain real world policy changes in regards to clear cases of individual moral responsibility for harmful outcomes. With regards to these changes, I will focus on how the moral responsibility of the individual citizen affects our intuitions in the just war scenario. I will ultimately conclude that moral responsibility can make the non-combatant citizen morally liable to deliberate attack under certain circumstances, because they are morally responsible for the ultimate policy outcome of the unjust war itself, through their voting action.

The overall strategy is to first apply a direct account of individualist moral responsibility. This is an account of moral responsibility which grounds the moral responsibility of the individual in her direct contribution to the relevant outcome. I show that while such an account can incriminate individuals in certain ideal small scale scenarios, the prospects for doing so in ordinary large scale ones, are poor. This is due to a failure to fulfil the certain causal and cognitive conditions which are usually associated with direct moral responsibility. Specifically, it is difficult to

account for how the individual can make a sufficiently significant causal contribution to the outcome, and further, the individual will usually be excused for believing that her contribution will be insignificant. In light of this, I examine whether an indirect account of moral responsibility (specifically shared responsibility) can justify holding the citizen morally responsible. I focus particularly on Christopher Kutz' complicity account, since this is as far as I am aware the most comprehensive, as well as the least demanding account of this sort. In the end I conclude that though promising, it is still too demanding with regards to the conditions for when something is the relevant sort of joint action, and further, it lacks a clear account of 'contribution' which allows for successful application to the voting case. Building on his account, and conclusions drawn throughout the project, I then construct an account of "moral membership", which can ground holding the individual citizen morally responsible for policy outcomes. The main developments here is that it allows for individuals who have clearly adversarial motives, to incriminate themselves in the same group project. Further, it involves a novel notion of "contribution", where the individual qualifies for moral membership if it was – assuming fulfilment of other background conditions – a reasonable expectation that she would be an interdependent part of the project that brings about the policy outcome. As noted, I tie this account to the cases of non-combatant citizen liability in the last chapter.

1.2 Chapter Outline

Chapter 1 - Introduction

In this chapter I introduce and elaborate the overall question, i.e. "Is the individual citizen morally responsible for policy outcomes?" After this I motivate why this is an interesting question by giving an account of what is at stake in answering it. In doing so, I put some perspective on how answering this question ties into the tangential debate of the liability of citizens in war. I also give an account of the real world relevance and possible implications of this work, which relate to how this project could either affect the behaviour of citizens, and/or help to tear down certain illusions the citizen may have with regards to her democratic idealism. Specifically, if it can be shown that the individual citizen is morally responsible, this will presumably justify an increased focus on political participation by the

individual. On the other hand, a negative answer could potentially justify a shift in focus away from the political participation of the individual, and to an extent undermine the normative status of democracy as an ideal worthy of pursuit. Lastly, I briefly document the conceptual room for this project, and note how little attention this particular question has received. At the end of this chapter, the groundwork for the rest of the project should be clear.

Chapter 2 - Direct individualist moral responsibility

In this chapter, I introduce the distinction between 'indirect' and 'direct' accounts of moral responsibility. I then focus on direct accounts, and I then disambiguate the term to distinguish it from other notions of responsibility. The goal is to bring out the particular concept of moral responsibility that I will apply more clearly. Part of the reason for doing this, is to narrow the discussion to come in the following chapters, and the other part is to avoid our intuitions running together due to confusing various distinct concepts. After this, I examine two different substantive interpretations of moral responsibility. Then I give an account of the particular theory I will apply to the overall question in this, and the next two chapters. Further, I examine and discuss the different conditions the individual plausibly has to fulfil in order to be morally responsible on a direct account. In light of this, I propose a strategy for testing whether the individual citizen fulfils the conditions for moral responsibility. After having done this, it should be clear how an account of moral responsibility can be applied to cases such as the current one, and the parameters for successfully answering the overall question should be set.

Chapter 3 - Causal condition

In this chapter I document that causal responsibility - the notion that a person or a thing actually causes the relevant outcome – seems to be a necessary condition for direct moral responsibility for that outcome, while highlighting certain problems with this idea. I then show that the voting case constitutes a severe problem for causal responsibility. The reason it constitutes such a problem is due to the particular threshold nature of the voting case, which means that even if the individual votes for the relevant option which would bring about the relevant

policy outcome, and even if that outcome wins out, the number of voters who voted for that outcome, overdetermined the outcome, i.e. were more than enough to bring about the outcome, in a way where it is difficult to say that the outcome depended on the contribution of the individual voter. I then propose a solution to this problem, but argue that this solution leads to counter-intuitive results. I then conclude, that if direct moral responsibility entails causal responsibility, then this undermines the conclusion that the individual is directly morally responsible for policy outcomes.

Chapter 4 - Cognitive condition

In this chapter, I start off with the assumption that the individual is either causally responsible for policy outcomes, or that moral responsibility does not entail causal responsibility. As argued in chapter 3), both of these assumptions are highly contentious. However, making these assumptions allows me to examine the second central condition to moral responsibility, namely the cognitive condition. Another reason for examining this condition is that it is plausibly also relevant to an indirect account of moral responsibility, which I will examine in later chapters, and which does not presuppose that the individual makes a direct contribution to the outcome. In light of these assumptions, I then examine whether the individual citizen fulfils this condition. Specifically, I examine whether she is aware that she makes a contribution to the outcome, and whether she is aware of the significance of the outcome itself. Further, I examine if her potential ignorance is excusable. In the end, I conclude that the individual citizen plausibly lacks sufficient awareness in the above sense, and that she is indeed excused for this. Most importantly, though she may be aware of the significance of the consequences of of a given candidate winning, she is not aware that she makes a morally relevant direct contribution to the voting outcome, even if she actually happens to do so. Further, she is excused for her ignorance, even in cases where she actually makes a direct contribution.

Chapter 5 - Rounding off all the previous chapters on direct moral responsibility

This chapters sums up the last three chapters in anticipation of the latter part of the project which focuses on giving an account of indirect moral responsibility for policy outcomes.

Chapter 6 - Indirect accounts of moral responsibility

In this chapter and the next, I examine an account of responsibility which I argue belongs under the heading of 'indirect moral responsibility', and see whether it can ground moral responsibility for policy outcomes. The approach is to examine a varied collection of theories, and then single out the most fruitful approach to focus on in the rest of the project. This is continued in chapter 7. Examining these theories will be helpful in further establishing the requirements of a successful account of moral responsibility for policy outcomes.

Chapter 7 - Joint action and complicity

As noted, this chapter is a continuation of the previous. Here I single out a *joint action* account of indirect moral responsibility as the most promising type of account so far for providing a positive answer to the overall thesis question. I focus particularly on Christopher Kutz' complicity account, since this is arguably the most comprehensive as well as the least demanding account of this sort. In the end I conclude that though promising, it is still too demanding with regards to the conditions for when something is the relevant sort of joint action, and further, it lacks a clear account of contribution which allows for successful application with regards to the voting case.

Chapter 8 – Moral membership

In this chapter I build on all the work in the previous chapters and tie it together to establish an indirect account of moral responsibility which I call a 'moral membership' account. As I argue, this account can successfully ground moral responsibility in the type of cases which are the focus of this project. The most important developments in this account are that though it is quite similar to Kutz' complicity account, it does not rely on the collective project being a joint action in a strong sense, if in any sense at all. Beyond that, it incorporates a notion of 'contribution', where an agent contributes to the particular project that brings

about the policy outcome, if her action helps constitute and sustain that project. When she contributes to this project, she is morally responsible for the outcome of the project. Specifically, I will argue that she contributes to the project when her contribution is an interdependent part of that project. She is an interdependent part of that project, if she is likely to be part of the causally efficacious set of votes (or simply 'causes', if it is not a voting scenario specifically) that brings about the voting outcome, and thus ultimately the policy outcome. If she is an interdependent part of the project, she contributes to the group project in a way which makes her a moral member of that project, even if she fails to make any direct contribution to the policy-outcome itself. Again, if she is a moral member of that project, she is morally responsible for the outcome of that project also. As I will argue, this account of indirect moral responsibility can be successfully applied to the voting case, which supports a positive answer to the overall project question.

Chapter 9 - Conclusion

In this chapter, I sum up and draw conclusions on the whole project. I apply the moral membership account to the voting case, and tease out the implications of the individual's moral responsibility. In doing this I also show how the conclusion from this project ties in to another discussion tangential to this one, namely the discussion of non-combatant liability in a just war scenario. Here I argue that the citizen's liability to being attacked in a just war, can be justified if the citizen is morally responsible in light of her moral membership of the project that brings about the harmful policy outcome, at least, in certain specific circumstances. This suggests that moral membership in a project that brings about a policy outcome can justify holding severe sanctions in light of her moral responsibility.

1.3 Unpacking the question

In this section, I will unpack the overall question: "Are Individual Citizens Morally Responsible for Policy Outcomes?". This will help in narrowing and clarifying what this project is about, highlight the broadest distinctions, and help specify what is at stake in answering the overall question. I will postpone introducing specific theory to the next chapters.

The first thing to highlight is that the focus of this project is the '**individual**'. This is meant to distinguish this project from other projects which focus e.g. on corporations¹, institutions and social structures² as the causes of outcomes, and the bearers of responsibility for outcomes. The key actor, and the relevant level of abstraction for this project, is the ordinary atomistic individual human citizen who forms (or at least revises) intentions in light of reasons, acts on these, and brings about events in the world in the everyday sense of performing actions. When a female secretary in a large oil firm destroys important documents which would have constituted legal evidence of companywide negligence, this may in part be described as *the company* acting as a corporate agent, performing harmful acts. Further, the fact that the female secretary is compelled to “shut up” and toe the company line, may be explained by women’s subordinate position in the patriarchal social structure, in the company or society at large. But, even if these ways of describing the situation are appropriate, it is plausibly also appropriate to understand her actions and responsibility as pertaining to her as an atomistic individual rational agent, reflecting on salient aspects of her situation, and responding to them. She makes a decision to destroy the documents, and depending on her circumstances, she may be morally responsible for doing so, and for the outcomes that follow from that action. This last understanding is what is indicated by “individual” in the main question.

The term '**Citizen**' in the overall question, highlights that the particular individual agent in question is the ordinary citizen. It is not the politician who is capable of making a direct impact on policy decisions. It is not the billionaire who affects

¹ I.e. ascriptions of responsibility to collectives as collectives. See e.g. French, 1979 and Pettit, 2007.

² The notion of ‘social structures’ has more common usage in sociology than philosophy, and though there seems to be various understandings of what social structures are, a central feature is that they are an organisational relationship between individual organisms in a society in a “complex network of social relations” (Radcliffe-Brown, 1940). These structures are social phenomena, and not obviously the product of any individual’s deliberate control. Therefore, though these can presumably explain why citizens bring about certain policy outcomes, they are not the pitched at the level of the individual agent in the individualistic sense I have in mind.

policy through lobbying. It is not the military commander who can single-handedly bring about large scale morally significant outcomes by giving a single order. It is the average citizen, the “you and me”, with ordinary capacities, and means of affecting policy and political actors. Often when we discuss politics or political issues, especially when something has gone awry, we tend to focus on key actors and what they did wrong, and what they ought to have done instead. So it is absolutely clear that those concerns occupy us. However, since the focus here is on the ordinary citizen, it will be necessary to tie the individual citizen to those key actors or the policy outcomes themselves. Then, depending on how those actors and policy outcomes are tied to the individuals, this may ground their moral responsibility for those outcomes. More specifically, I will be focusing on the individual citizen residing in what I will refer to as an ‘imperfect democracy’. This is inspired by Walzer (1977), who contrasts this to a “perfect democracy”. A perfect democracy is one which perfectly instantiates the will of the people, where citizens come together, in the strain of deliberative democracy and rationally discuss political issues. They have a great deal of insight, openness and willingness to understand the various issues, and they then together work for the optimal solution to the relevant challenges, and then bring about these solutions through the voting act. ‘Imperfect democracy’ is labelled as such, because democracy usually falls short of this ideal. The public is not always well informed, and often too ideologically entrenched to allow for open unconstrained reflection on the relevant issues. Further, implementation is made problematic by lobbyists and various political interest groups that try to affect policy in light of their own narrow agenda. Though imperfect, it is still appropriate to assume that the citizenry has a strong influence over the policies of the state, at least by way of voting. In light of this, I will for the sake of this project be imagining something in the order of a society with fair and equal elections among free adult citizens, whose behaviour is in relevant aspects only constrained by certain important constitutional limitations. I will assume this is an appropriate model of democratic society. I imagine that this society includes common western liberal provisions such as at least a set of basic liberties. The reason for assuming this type of society is due to the increase in relevance of a discussion of citizen responsibility, in a

reasonably open society with as many *de facto*³ freedoms and degrees of influence, such as those common in “western” liberal democracies. In societies where the individual has rather limited (or none at all) freedoms, or power by which she can hope to influence policy outcomes, the discussion of individual responsibility seems fairly trivial. In authoritarian states there is no clear tension between the will of people vs. the responsibility of the individual. The reason is that it is not clear that individual citizens have any significant channels of influence. Individual influence is curtailed by the ruling elite and expressions of political participation are at best fiercely manipulated and at worst entirely controlled.

The term '**Moral responsibility**' in the main question requires a significant amount of unpacking itself, and I have reserved a number of chapters to doing just that. In broad terms, the way I will understand the term here, moral responsibility refers to when some action or outcome can appropriately be attributed to an agent as something he or she *has* done or brought about retrospectively, which in morally interesting cases at the very least merits an ascription of praise and blame for that action or outcome, and in some cases further sanctions e.g. in the form of punishment and/or compensation. A straightforward quick and dirty example of this notion of moral responsibility applied, is when a person, without clear previous provocation, deliberately, and freely harms another person with the aim of e.g. making her suffer. In such a case, we clearly understand the perpetrator as being morally responsible for the harm in question. Further, we would presume blame her for harming the other person. Lastly, if the amount of harm is above some threshold, we would at least consider it justifiable to put the perpetrator in jail. Though this example is clear enough, the terms ‘responsibility’ and ‘moral responsibility’ are quite ambiguous. Therefore I will devote a large amount of space to elaborating the concepts I will be focusing on, and to distinguish them from related notions of responsibility. With regards to moral responsibility, I examine the citizen’s moral responsibility in light of two quite distinct categories of moral responsibility, namely ‘direct’ and ‘indirect’ accounts.

³ By ‘*de facto* freedoms’ I am simply contrasting liberal democratic societies with societies where the individual may certainly have constitutional rights, but where these are entirely formal.

Regarding a ‘direct account’⁴, this concerns whether the relevant outcome can be attributed to the individual agent as something she produces or contributes to directly, by herself, where the outcome depends to some extent on her making a significant contribution to the outcome itself. As I will show, an ascription of moral responsibility on a direct account depends at least on two necessary conditions, namely a ‘causal’ and a ‘cognitive’ condition. The causal condition (also referred to as causal responsibility) concerns whether the outcome obtaining directly depends on the contribution of the individual. The cognitive condition concerns whether the individual is aware of her contribution, or ought to be aware of her contribution, when she performs the relevant action. It also concerns whether she is aware of what the actual outcome entails. If the individual does actually make a direct contribution to the outcome, and actually has good reason to think she does, or at least ought to see herself as making a contribution, then she is – assuming she fulfils other background conditions which I will elaborate in the subsequent chapters⁵ – morally responsible for policy outcomes. Regarding indirect accounts, this is a broad category of accounts which all have in common that they try to explain how the individual can incriminate herself in an outcome even when she fails to make a direct contribution to the outcome itself. I will focus mainly on what can be broadly construed as accounts of ‘shared responsibility’. On these account, the individual can be said to share in the moral responsibility of an outcome she and others bring about, again, even if the individual fails to make a significant contribution to the outcome itself. I will expand on the direct account in the next chapter, and the indirect accounts from chapter 6 and onwards. As I will show, the indirect accounts are more promising

⁴ A ‘direct’ account of moral responsibility is also often referred to as a ‘control account’, but I prefer “direct” to avoid confusing it with indirect accounts, since indirect accounts may also involve some aspect of control. The difference between the two categories concerns mainly what it is we exercise control *over*. On direct accounts, it is the outcome itself. On indirect accounts, it is usually not the outcome itself.

⁵ I will elaborate all the conditions in chapter 2, but here note that they include, but are not limited to: that the individual is not being manipulated, that the individual is sensitive to moral concerns, that the cost of defection is sufficiently low (e.g. no coercion), and fulfilment of some freedom relevant condition, e.g. having compatibilist or libertarian free will.

than the direct one in terms of giving an account of the moral responsibility of the individual for policy outcomes.

'Policy outcome' indicates that the focus of the project is not simply on whether the individual citizen is morally responsible for an action, or a "trying" to bring about an outcome. It concerns an outcome that actually comes about, and whether that policy outcome is then something which can be tied to her as something she brought it about, or something she made a significant contribution to. A policy outcome is here understood as the ultimate consequence of e.g. a government enacting a certain policy, or of a military general ordering an attack, or some further mechanism for bringing about an outcome. The key distinction here is that it is the actual consequence of a policy, not the proposed policy in itself, which the individual is morally responsible for. E.g. the policy may be declaring war against a neighbouring country, but the policy outcome is people dying and the destruction of infrastructure. This is a relevant distinction because the ultimate consequence may be reasonably unforeseen, making it more difficult to attribute it to the individual citizen, even if she is making a contribution to that outcome. We can imagine someone who votes for Hitler, and Hitler wins the election, in part due to that person's contribution. But if there was no way of anticipating the Holocaust, e.g. because it was not in the plans at that time, or because the public was misled, it is not obvious that this person is then morally responsible for that ultimate horrific outcome, which is what is of relevance here. This pertains to the cognitive condition for moral responsibility which I will discuss in chapter 4. On the other hand, there are presumably also cases where the citizen has a reasonable expectation about some harm being a consequence of some candidate winning, even if the candidate is never explicit about it. Voting for someone who is clearly a "loose cannon", should presumably be anticipated to have potentially harmful consequences down the line. Still, in this case it is also the ultimate somewhat foreseeable outcome that is relevant to the ascription of moral responsibility. Lastly, this project does not concern an ascription of moral responsibility for a given candidate, party or policy in isolation aside from the policy outcomes which are a result of their electoral success. Of course, putting an evil candidate in office may be a necessary condition for bringing about the evil

policy outcome, but the outcome itself is what is the central object of an ascription of moral responsibility with regards to this investigation.

Though not explicitly stated in the question, '**Political participation**' will as indicated, be considered the main method for tying the policy outcome to the individual. The specific exemplification of this in this project, will be the act of *voting*. There are presumably many acts of political participation which can ground the individual's moral responsibility. However, there are good reasons for focusing on voting in particular. As noted, I will mainly understand the morally relevant policy outcome as being tied to the individual in light of her political participation. The reason is, voting is arguably the clearest way in which the individual can express her will and make a contribution to policy outcomes. It is for most people considered the paramount example of political participation. It is unlikely that we will openly criticize someone who fails to write a convincing blog post, or fails to participate in a demonstration for or against some important issue, or fails to articulate her views at *speakers' corner*. But if someone declared that she does not want to vote, this is usually met by at least astonishment, followed by a list of reasons why it is very important for a citizen to vote. A second reason for this focus is simply a matter of research practicality. There has been a good deal of writing focused on voting specifically, both in philosophy and related fields. It is also easy to measure the impact of voting compared to other forms of participation, which makes it more straightforward to bring out certain important considerations pertaining to individual responsibility, such as the individual harm caused, and the foreseeability of causing harm. It is of course important to examine whether any findings can be generalised to other forms of political participation, but it is outside the scope of this project. Presumably however, the developments of this project can be generalized to other forms of political participation.

Summing up: the overall question concerns whether the ordinary individual citizen through her contribution, can be morally responsible for the morally significant policy outcomes of the state, to an extent which merits an ascription of praise and blame, and which would potentially justify certain sanctions. To return

to the hypothetical historical example above: it e.g. concerns whether the German citizen by voting for the NSDAP, while having a reasonable belief that if this party wins they will enact a policy of genocide, which will lead to great amounts of preventable death and destruction, is morally responsible for this death and destruction, or part of it, when it actually occurs, either because she brings about the outcome directly, or because she otherwise is in some morally relevant way indirectly associated with it.

1.4 Why it is an interesting question

In this section, I will note what the practical implications of this work may be, and highlight how this work relates to the philosophical landscape.

There are several reasons for why the overall question is interesting. Firstly, if a large proportion of the public is indeed responsible for policy outcomes, the fact that they are not perceived as appropriate bearers of responsibility may be wrong in its own right. Furthermore it may have unfortunate consequences. E.g. if citizens fail to regard political participation with the appropriate severity it deserves, this could lead citizens to act carelessly when executing political power. If citizens do not take their responsibility seriously, there is a chance they will make political decisions on a whim, or allow themselves to be directed mainly by their immediate passions, possibly leading to suboptimal and simply harmful policy outcomes. They may vote for or otherwise support morally bankrupt political candidates, or they may support candidates without taking enough time to deliberate on the consequences of these candidates winning. Further, the wrongs of the state are in many cases passed on to the citizens of the state. E.g. when a state is sanctioned for breaking certain international conventions, the burden will almost surely be passed on to the citizens of the state, e.g. through increased taxation, or reductions in social goods. Giving an account of individual responsibility may at least in part explain and help justify imposing such sanctions on the relevant morally responsible citizens. This is particularly relevant when questions of citizen liability in war are raised, where this may depend on answering the question of their moral responsibility for the unjust war that has been brought about in part due to their political participation.

On the other hand, a negative answer to the overall question may support the notion that we are usually actually fully justified in a casual stance towards our political support and politics as a whole. If this is the case, then a conclusion from this investigation may support a critique of the notion of democracy “representing the will of the people”, which many westerners see as something making this particular system comparatively superior or ideal. Further, it may support the case that citizens are not liable to attack in war. Therefore, it will be beneficial to clarify what is at stake in terms of policy outcomes, and determine whether individual citizens are indeed morally responsible for these outcomes. As noted, many, if not most people have an interest in political matters. Unless this is simply a pastime interest, it seems odd that people would care for these matters, if they acknowledge that they are completely detached from them and unable to do anything to affect whether the policy outcomes actually comes about. Indeed it seems plausible that one reason people are interested in politics and policy outcomes, is because they actually believe they can impact them, at least in some sense. This project is then an attempt to examine how the individual can impact policy outcomes, or can otherwise be tied to them, and to examine whether this relationship is one that can ground moral responsibility for policy outcomes. As noted, I will ultimately argue that it is appropriate to hold the individual citizen morally responsible for policy outcomes.

Regarding the philosophical landscape, this project examines the moral responsibility of the individual, in light of two fairly distinct approaches to moral responsibility, namely ‘direct’ and ‘indirect’ approaches. A direct approach attributes an outcome to the individual in light of her direct contribution to the outcome. In terms of direct moral responsibility, I am unaware of anyone who has tried to justify an ascription of moral responsibility for policy outcomes. I imagine that this is because it is considered an obviously futile effort. Indeed, on the face of it, that certainly seems to be the case. This is presumably because it is implausible that the individual can make a sufficiently strong contribution to the voting outcome, for this policy outcome to be attributed to her. Still, though I will eventually agree, it is important to examine precisely where such a project breaks

down. This will help both motivate an indirect account, as well as highlight those challenges an indirect account needs to respond to, in order to give a positive verdict to the overall question. Regarding 'indirect' approaches to moral responsibility, the core notion here is that the individual can be morally responsible for an outcome beyond her direct contribution to that outcome. However, it is still distinct individuals acting as individuals who are the main agents for an ascription of moral responsibility. I will examine and discuss such approaches from chapter 6 and onwards. As perhaps expected, there has been some attempts to try to account for the moral responsibility of the individuals on this approach. However, the central question of this thesis remains more or less untouched, with a few notable recent exceptions which I will examine in more detail in later chapters (Pasternak, 2011, 2012, Abdel-Nour 2003, May 1992, Kutz 2000, Seumas Miller 2001, Gregory Mellema 2006). These accounts are quite varied in how they justify an ascription of moral responsibility. Some⁶ argue that individuals may stand in a client/agent relationship with the state, where they are in a position to authorize the state and the harms that states bring about, where doing this confers responsibility on the individual. Others⁷ focus on how the citizens identify with the outcomes of the nation, to an extent which can also make us responsible for these outcomes. Further, and as I will argue, most promising, others⁸ focus on whether what we do as citizens can be construed as a responsibility-grounding joint action. The particular account I will construct is most clearly related to this last category of approaches. Such accounts try to ground the individual's moral responsibility for a given outcome, in her participation in what can appropriately be characterized as a joint action, where the morally relevant outcome is then brought about through this joint action. These accounts are distinct from direct accounts because it is usually not necessary that the individual herself makes a direct contribution to the outcome in order for her to at least share in the moral responsibility for that outcome, as

⁶ E.g. Pasternak (2011, 2012), Parrish (2009), Young (2004).

⁷ E.g. Abdel-Nour. (2003), May (1991) and Jaspers (2000).

⁸ In this project, Kutz (2000) specifically. But relevant to this view are e.g. also Gilbert (2006), Tuomela (2010), Bratman (1992).

long as she can appropriately be said to be an inclusive part of the joint action. Though my account is clearly reminiscent of a joint action account, it is not an account of joint action proper. It is e.g. less demanding with regards to the cooperative attitudes of the participants, allowing decidedly adversarial attitudes among the participants, to the extent that they can incriminate themselves in each other's projects even if they try to sabotage each other, or would if they could. Further, I give a novel account of how the individual can be said to contribute to the joint project, where we contribute to a joint project if it was likely that we were a foreseeably interdependent part of the set of causes that actually brought about the project. I incorporate this notion of contribution into an account of indirect moral responsibility, which is applicable to the voting case, and in light of which we can hold the individual citizen morally responsible for policy outcomes.

Lastly, though there has not been a great amount of work done on individual moral responsibility for policy outcomes specifically, there is relevant tangential work which connects to my project. Particularly one group of views are especially relevant for this project, namely so called reductivist accounts of non-combatant liability. These views are relevant to this project, because they hold that the criterion for individual liability in war, may depend on the moral responsibility of the individual. (e.g. McMahan 2009, and Primoratz 2005). Here, one view suggests that non-combatant liability to attack in war can potentially be grounded in the non-combatants moral responsibility for the harm which justified military intervention in the first place, if this attack can be justified as *defensive killing*. After having produced a successful account of moral responsibility for policy outcomes, I will apply this account to the just war scenario. This will allow me to show that moral responsibility potentially grounds quite severe moral evaluations, to an extent where it may be appropriate to attack the citizen because of her voting act. This then adds to the discussion of non-combatant liability by supporting an important part of the argument for non-combatant liability, i.e. the "moral responsibility" part.

1.5 Summing up

In the following I will approach the overall question by resolving the tension documented earlier between notion that the will of the people is represented in the policy outcomes of the state, and our apparent reluctance of holding our fellow citizens morally responsible for these outcomes. I will do this by examining how we can justify holding the individual citizen morally responsible for policy outcomes. Specifically, I will firstly examine a direct control account of moral responsibility, which I will show fail in this regard, because the individual fails to make a significant direct contribution to the policy outcome, and would even be excused for doing to if she did. Then I use this failure to motivate looking towards indirect accounts of responsibility, which I review. Even though they are unsuccessful, I will build on them, and ultimately construct an account of “moral membership”, which can be successfully applied to the overall question. After this, I will present the conclusions of this project, and draw some further implications of these conclusions, specifically with regards to the citizen’s (non-combatants) liability to being attacked in war. I will show how the developed account can in some cases ground the defensive killing of civilians.

2 Direct individualist moral responsibility

2.1 Introduction

In this chapter, I introduce and examine what roughly falls under the heading of a ‘direct’ account of moral responsibility.⁹ The aim in this chapter is firstly to specify and elaborate the particular broad concept of responsibility I will apply to the voting case in the first part of this project (chapter 2-5). The second aim is to develop the particular strategy I will employ in order to test whether the individual is directly morally responsible. In order to reach these two aims, I will firstly present this responsibility concept, and distinguish it from other responsibility concepts. After this, I will examine how we ought to interpret this concept in a more substantive sense with regards to what it means to *be* morally

⁹ I say roughly, because there is an overlap between direct and indirect moral responsibility.

responsible, and what it means to *hold* someone morally responsible. This will be important in order to clarify the specific conditions for moral responsibility, and attaining the second aim of developing the noted strategy. At the end of this chapter, I will have set out a clear strategy for the next two chapters where I test whether the individual citizen is *directly* morally responsible for policy outcomes.

The structure of the chapter is as follows: I will firstly introduce and examine the notion of direct individualist moral responsibility I will discuss and apply in the subsequent two chapters. Specifically, in section 2.2 I distinguish moral responsibility from other concepts of ‘responsibility’ in order to single out the particular concept of interest to this project. Then in 2.3 I elaborate the relevant notion of *moral responsibility* in greater detail. Here I highlight two common ways of understanding this notion, namely in accordance with a Strawsonian account of moral responsibility, and in light of the so called “ledger view” of moral responsibility. After this, I narrow my approach even more in section 2.4, where I focus on a so called ‘control’ account of moral responsibility in order to derive the specific necessary and sufficient conditions for moral responsibility. In the last part of this section I set out the strategy for testing these conditions which I will do in the next two chapters. I will sum up in chapter 5.

2.2 Responsibility concepts

The term ‘responsibility’ is quite ambiguous. HLA Hart has e.g. famously identified at least four distinct notions of responsibility (Hart 1968:211 and Tognazzini 2013), and other writers have identified as many as six (Vincent 2011). Therefore, it will be important to provide some clarification of the particular broad notion I will be discussing and applying to this project. I will present Hart’s parable of the drunken sea captain, and use it as a backdrop for clarifying the broad notion of direct retrospective moral responsibility. I will not give a complete account of all responsibility concepts in the parable, since they are quite well known, and not all relevant to this project.

Drunken sea captain:

“As captain of the ship, X was responsible for the safety of his passengers and crew. But on his last voyage he got drunk every night and was responsible for the loss of the ship with all aboard. It was rumoured that he was insane, but the doctors considered that he was responsible for his actions. Throughout the voyage he behaved quite irresponsibly, and various incidents in his career showed that he was not a responsible person. He always maintained that the exceptional winter storms were responsible for the loss of the ship, but in the legal proceedings brought against him he was found criminally responsible for his negligent conduct, and in separate civil proceedings he was held legally responsible for the loss of life and property. He is still alive and he is morally responsible for the deaths of many women and children.” (2008: 211).

‘Moral responsibility’, is – as indicated by its prominence in the thesis title – the main responsibility concept I will apply in this project. It is indicated by the last sentence of the parable, namely he is morally responsible for the death of the women and children. Moral responsibility is retrospective. This means that it concerns past events. Specifically those *actions* that *have* been performed, or *outcomes* that *have* been brought about through some direct contribution. Further, ‘direct’ moral responsibility concerns actions or outcomes the individual has brought about *as* an individual, as opposed to the actions or outcomes that have been performed or brought by a group. What makes a ‘direct’ contribution to an outcome *direct*, is that the individual is here exercising some strong *control* over that outcome, or part of it. I will elaborate what control means in a subsequent section, but note that it at least presupposes that the outcome was to some extent *causally* dependent on the action of the individual herself, and that she was aware (our ought to have been aware) of this dependence. ‘Indirect’ moral responsibility is *indirect*, because it usually does not rely on the individual exercising control over the relevant outcome. E.g. she can arguably (chapter 6 and onwards) be indirectly morally responsible for an outcome by merely participating in a collective project that brings about the outcome. As I will show, there are potentially various ways in which she can be indirectly morally responsible.

Moral responsibility concerns actions and outcomes which can be attributed to us as agents, which we are at fault for having produced, at least when these concern harmful events.¹⁰ When someone is (directly) morally responsible for an action or outcome, this entails that the action or outcome is something *they* individually have brought about in a way where the action or outcome is representative of who they are. Michael Zimmerman puts Moral responsibility under the broader heading of “personal responsibility” (in Clarke, McKenna, & Smith, 2015: 45) to indicate that this concept implies that the outcome is representative of who the person is, apart from merely being representative of her bodily movements. In order to be representative of who the person is, the person should also have been aware, or ought to have been aware, that he or she brought about the outcome. Further, for an ascription of moral responsibility, the action or outcome in question at the very least has to have some moral relevance, i.e. be harmful. Lastly, it is often assumed that if someone is morally responsible, some blame, and potentially punishment is appropriate. This is however a matter of contention. As I will note in the next section, this depends on the particular substantive interpretation one adheres to. Safe to say however, moral responsibility is often associated with our blaming practices, and other moral sanctions. When a person is also an appropriate target of some moral sanction, some will refer to the individual as being ‘accountability responsible’ (e.g. Watson 1996).

As noted, moral responsibility is retrospective. Moral responsibility can be contrasted to *prospective* responsibility concepts such as ‘virtue responsibility’ and ‘role responsibility’. Virtue responsibility usually implies what we mean when we refer to someone as a *responsible person*, as opposed to being an irresponsible person. Being virtue responsible then means that the person is *the type of person* who we can rely on prospectively to do the *right thing*, e.g. because their personal history can attest to this. It is indicated by “various incidents in his career showed that he was not a responsible person” in the parable. ‘Role responsibility’

¹⁰ I will ignore positive, or beneficial events entirely in this project with regards to retrospective moral responsibility, and thus the question of whether we can be morally responsible for positive outcomes. I will focus entirely on harmful actions and outcomes.

concerns prospective duties and obligations. It is indicated by “was responsible for the safety of his passengers and crew”. Though role responsibility is not retrospective, it can be derived from an ascription of retrospective moral responsibility. E.g. if I am morally responsible for dumping chemicals in the local water supply, I may in light of this, have role responsibility, i.e. an obligation to clean it up. Further, the fact that I have dumped these chemicals, may indicate that I am lacking in virtue responsibility. I am the kind of person who would do *such a thing*.

Summing up this short exercise: the individual is directly morally responsible for a harmful event that *has* obtained, when it depended on her action and she was aware of this dependence. In the next section and through this chapter, I will elaborate this in greater detail, and specify the individually necessary and jointly sufficient conditions. At the end of chapter 4, I will conclude that the individual fails to be morally responsible for policy outcomes on this account. In chapter 5, I sum up and round off these chapters of direct moral responsibility. This motivates moving on to chapter 6 and the subsequent chapters, to examine whether the individual is then morally responsible on an indirect account of moral responsibility.

Having distinguished *moral* responsibility from other responsibility concept, I will in the next section focus on the details of this notion of responsibility.

2.3 Substantive accounts of moral responsibility

Though ‘moral responsibility’ can be distinguished from other ‘responsibility’ concepts, as shown in the previous section, what the concept itself means in a more substantive sense, deserves further attention. This will be helpful in establishing the conditions for moral responsibility. I will firstly examine two different substantive interpretations of moral responsibility – the Strawsonian view, and the Ledger view. I will then contrast the two. Then I will introduce a plausible further distinction between attributability- and accountability responsibility, which will help in clearly distinguishing the specific necessary

conditions the individual has to fulfil in order to be directly morally responsible for policy outcomes.

Strawson, P.F. (2008) observes that when we acknowledge that someone produces certain benefits or harms, or that they act admirably or badly, we are often prone to certain “reactive attitudes” (2008: 6). These include a wide range of positive or negative attitudes, including but not limited to gratitude, shame, pride and indignation (Watson, in Schoeman, 1988: 257 for an extensive list). A central claim in Strawson's account is that these reactive attitudes are not just *associated* with moral responsibility, but are in fact *constitutive* of it. For Strawson, *holding* someone morally responsible in this way, is a fundamental part of being human, because it is fundamental to our interpersonal relationships. Doing away with such practices would presumably impoverish our lives to a great extent (ibid: 14). Further – barring certain pathologies – we cannot avoid having these attitudes, even if we wanted to.

Though our reactive attitudes are constitutive of moral responsibility, Strawson acknowledges that we can still to some extent *hold* other individuals morally responsible even if we lack the reactive attitudes. It may e.g. be appropriate to praise or blame someone for instrumental reasons, e.g. admonish them because we aim at affecting their behaviour prospectively. However, doing that involves having what Strawson calls an “objective attitude” (2008: 9), which means: “to see [that person], perhaps, as an object of social policy; as a subject for what, in a wide range of sense, might be called treatment; as something certainly to be taken account, perhaps precautionary account, of; to be managed or handled or cured or trained; perhaps simply to be avoided” (ibid). *Holding* responsible in this sense, is e.g. what we do in many cases when we praise or admonish a child for her good or bad behaviour in the hopes that it will have a positive effect prospectively. We act as if we have the corresponding reactive attitudes towards her, but such an action does not require the existence of genuine reactive attitudes. The objective attitude is special, and it is not something which is feasibly kept for extended periods of time. Even if it could, it would be a shallow experience. Susan Wolf puts this most poignantly when she states: “A world in

which human relationships are restricted to those that can be formed and supported in the absence of the reactive attitudes is a world of human isolation so cold and dreary that any but the most cynical must shudder at the idea of it" (Wolf, S. 1981: 391).

In order to understand why Strawson focuses on our reactive attitudes like this, one has to look at what he is responding to, namely the threat of determinism. If everything in the universe is determined, some believe this rules out the possibility of moral responsibility. One line of reasoning is that we cannot be responsible for something if we lack some degree of freedom of choice regarding whether the action or outcome we are purportedly responsible for comes about. If the universe is determined, and thus if all events are already fixed, that seemingly undercuts anyone's choice in the matter of how the events we are seemingly morally responsible for unfold, and whether they obtain or not. In that case, we can seemingly not tie these events to the individuals as *their* doing. Strawson's account is an attempt to circumvent this problem. Broadly speaking, his claim is that whether everything in the universe is determined or not, it will not constitute a reason for us to give up our reactive attitudes, and thus not our notion of moral responsibility (because that is what moral responsibility amounts to).

The argument builds upon the insight that there are two possible types of *pleas* with regards to moral responsibility, and the observation that the thesis of determinism fails to adhere to fit with any of them. A plea is a reason to modify our reactive attitudes, i.e. a reason to adjust our ascription of moral responsibility somehow. (McKenna 2012: 74) The first type of plea (excuses) concerns when something makes us readjust our reactive attitudes, e.g. in light of new information about the agent and/or her circumstances. I may be resentful of Jones because he kicks his dog, but if I come to know that it was just a cardboard cut-out of a dog, my attitudes will most likely change. The first type of plea concerns e.g. knowledge of the person's aim, or the information the person had access to, or other things which might lead us to acquire a different understanding of why the person acted as he did. If the person e.g. had good

reasons to think it was a cardboard cut-out, but it in fact was not, this will presumably affect our reactive attitudes, from when we thought Jones knew perfectly well that it was a real dog, i.e. we may excuse him, at least to a degree. The second type of plea (exemptions) is one that undermines the moral agent as an appropriate recipient of reactive attitudes altogether, either in the current situation, or entirely. I may again see Jones kicking his dog, but if I come to know that he's severely mentally handicapped, and unable to control his actions or unable to understand that dogs are not simply playthings, this may lead me to consider that he is not an apt candidate for reactive attitudes at all, that he's outside the scope of them. In this case, it is not new information about his reasons for action which affects my judgment. It is information about his overall moral status. The second type of plea undermines the agent's moral status, e.g. if the person has a mental illness, or is a child, or is under hypnotic suggestion (Strawson, P. 2008: 8). Whatever the case is, it makes them inappropriate targets of reactive attitudes. In this case, we often instead shift to the objective perspective, and relinquish our reactive attitudes, at least ideally. We at best consider them an *object* of treatment. Regarding the second type of plea, though it may in some cases be perfectly appropriate to take on the objective perspective, our attitudes here cannot "include the range of reactive feelings and attitudes which belong to involvement or participation with others in interpersonal human relationships; [they] cannot include resentment, gratitude, forgiveness, anger, or the sort of love which two adults can sometimes be said to feel reciprocally, for each other" (Strawson, P. 2008: 10). The reason is, this person's ailment has diminished our ability to have and maintain a normal human relationship with that person.

The question is then whether determinism constitutes either type of plea. Strawson's claim is that the thesis of determinism does not undermine moral responsibility, because it does not fit into either type of possible of plea, i.e. excuses or exemptions. Regarding the first type of plea – excuses – they only apply to particular local actions, and their application rests on the assumption that moral responsibility is possible. Since determinism is a global thesis that applies to all phenomena all the time, it does not apply to these particular

instances. Determinism is always in play, so it cannot constitute an excuse. If determinism is moral responsibility undermining, it must pertain to the second type of plea. However, the claim that the acceptance of determinism would constitute an exemption to moral responsibility all together is – according to Strawson – implausible. The reason is that the acceptance of determinism will not lead us to give up our reactive attitudes, and since they are constitutive of what moral responsibility is, determinism will not make us give up moral responsibility. Now, we may simply disagree with him and object that descriptively, if we really did accept the thesis of determinism wholeheartedly, we in fact would give up our reactive attitudes. Strawson tries to convince us of the opposite when he writes: “I suppose we must say that it is not absolutely inconceivable that it should happen. But I am strongly inclined to think that it is, for us as we are, practically inconceivable” (2008: 13). The reason is that he believes our praising and blaming practices are: “too thoroughgoing and deeply rooted for us to take seriously the thought that a general theoretical conviction might so change our world that, in it, there were no longer any such things as inter-personal relationships as we normally understand them; and being involved in inter-personal relationships as we normally understand them precisely is being exposed to the range of reactive attitudes and feelings that is in question” (ibid). It is of course an empirical question whether this is true or not, and I will not try to challenge it here.¹¹ The further question is then whether we *should* give up our reactive attitudes in light of acceptance of the determinist thesis. With regards to this, Strawson claims that “the truth or falsity of a general thesis of determinism would not bear on the rationality of this choice” (2008: 14). Rather, when tallying up the pros and cons of whether we should give up our reactive attitudes, it is rather quality of life considerations which are central to such a judgement, not the thesis of determinism at all, i.e. it is whether it would make our lives better or worse. But the fact that it is a reasonable question to ask, and the fact that an answer cannot hinge on the truth or falsity of the thesis of determinism, indicates again that determinism is not any relevant concern with regards to whether moral

¹¹ This is exactly what Knobe, J. and Nichols, S. do in their 2007 paper, where they show empirical data which may suggest that these intuitions can be challenged.

responsibility is undermined by determinism.¹² So Strawson leads us to the conclusion that determinism will not, and ought not undermine our reactive attitudes, and thus not our practices of holding each other morally responsible, because it does not fit with the second plea either. Because it will not undermine our reactive attitudes, it does not undermine the thing that is constituted by them, i.e. moral responsibility. Thus, determinism does not constitute an exemption. In that sense, his position is compatibilist. Instead of giving a justification for why we should hold someone responsible, Strawson tries to show us that it is simply an inescapable part of human practices, and in practice, it does not hinge on the truth of determinism at all.

Strawson's view has been extremely influential, and has led to a naturalistic turn in the debate of moral responsibility (Kane, R 2012: 200). Focus has shifted from the abstract metaphysical question of enquiring what a morally responsible agent *is*, to instead enquiring into what is actually involved in the practice of *holding* someone morally responsible, and the moral psychology of examining the associated constitutive attitudes. A point of contention here is that if our moral responsibility is simply the practice of *holding* responsible, this seems to run contrary to the observation, that in many cases, when we praise and blame someone, it is still an open question whether they *are* morally responsible. Specifically, it seems clear that our judgements can be wrong, because our psychological dispositions are malleable. Fischer and Ravizza e.g. point out that if a society in which the “severely retarded or mentally disturbed individuals are resented, blamed and harshly punished for their failure to adhere to the norms of the community” (1993: 18), it seems that we can still appropriately criticise these practices for being wrong. According to McKenna however (2008: 118) this conclusion is however too hasty. Our reactive attitudes still have an internal criterion regarding their appropriateness, so if we e.g. wrongly judge another person to be rude, e.g. because we are prone to hasty judgements from too much caffeine and too little sleep, and in light of this judgement feel resentful towards

¹² Interestingly enough, his son, Galen Strawson disagrees with this and claims that when we have reactive attitudes, and make judgments about this person being morally responsible, we also have a built in assumption that the person in question is not determined (Strawson, G. 2008: 88).

the person in question, this is actually a misapplication of our resentment, according to the internal standard of the attitude itself. We felt resentful, but inappropriately so. The person may be morally responsible per se, but the person will also have a valid excuse. Though our reactive attitudes are constitutive of moral responsibility, and moral responsibility here is nothing more than our reactive attitudes, according to the noted interpretation, this does not entail that our reactive attitudes are always appropriately applied, as indicated by the first type of plea. I may be compelled to form a reactive attitude because I think my friend is stealing from me. But when I see that I was mistaken in my judgement, e.g. because I come to understand that it was actually his own thing, I naturally adjust this attitude because it becomes clear that it did not correspond with the facts as they really are. Therefore, his account still allows for specifying the relevant conditions.

A different substantive view of moral responsibility, and closely related to the Strawsonian view, is the so called “ledger view” (e.g. Fischer & Ravizza, M. 1998, Zimmerman 1988, 2002). Those who favour this view presumably agree that when we hold someone morally responsible, this implies that we are usually ready to treat them in certain ways. But they either disagree with Strawson that reactive attitudes *constitute* moral responsibility, or they believe that we should loosen the connection between reactive attitudes and moral responsibility. They suggest that it is appropriate to say that someone is morally responsible as a factual claim, even if this does not thereby make a reactive attitude appropriate, or even if the agent is undeserving of being the target of any positive or negative attitude or action (e.g. retribution or reward). On the ledger view, an ascription of moral responsibility is logically independent of our reactive attitudes. On this view if we consider someone morally responsible, this implies merely their moral value has been affected (Fischer, J.M. & Ravizza, M. 1998). However, with regards to desert it is silent, and any *holding* responsible requires further justification. Any reactive attitude or moral sanction such as punishment depends on reasons aside from the ascription of moral responsibility. As Joel Feinberg writes: “To be morally responsible [...] is not to be liable to any kind of official action or even to unofficial informal responses such as acts of blaming” (1970: 30), setting it in stark contrast

to the Strawsonian view. The direction is the other way around than for Strawson (Watson, G. 2004: 222) with regards to our attitudes. On the ledger view, a common way of thinking about the issue, is to imagine a *metaphorical* moral ledger “or some ideal record, liability to credit or blame” (Feinberg, J. 1970: 30). When a person then is morally responsible, this means that they fulfil the conditions – whatever they are – to get a mark on their metaphorical ledger “It is as if, when a morally responsible agent does something praiseworthy or blameworthy, a new truth enters the scene, and there is a new “mark” on the moral “ledger” of that person’s life” (McKenna 2012:43). However, it is an open question whether reactive attitudes are appropriate. Moral responsibility is strictly an intellectually based assertion, rather than an interpersonal experience, as in Strawson’s case.

Comparing the two substantive interpretations, I do not think there are any clear necessary distinctions with regards to answering the overall question whether one takes the Ledger view or the Strawsonian view. On the Strawsonian view, it is an open question whether our reactive attitudes (and practices) correspond to the particular reactive attitude’s criterion. On the ledger view, it is an open question whether a given reactive attitude is justified. Supporting the notion that they are comparable in their application, Fischer and Ravizza (1998: 9-10) argue that one could indeed envision a mixed account of these two substantive views. On such an account, someone is an apt candidate for reactive attitudes, iff one has a positive or negative mark on one’s moral ledger. The ledger view then specifies the internal criterion for our reactive attitudes which matches the conditions for an ascription of moral responsibility.

Whether one subscribes to the Strawsonian view, or the Ledger view above, a further question concerns what the actual conditions for moral responsibility are (or what the internal criteria for our reactive attitudes is). In relation to this, a development after Strawson highlights that Strawson and others before him seem to conflate two distinct notions of responsibility. One is responsibility as ‘attributability’ and the other is responsibility as ‘accountability (Watson, G. 2004: 265). Attributability responsibility does not in itself ground or justify the actual act

of *holding* someone responsible, i.e. it does not in itself make praise, blame or punishment appropriate. It can simply be an insulated judgement about the person character, or a judgement about what the person did, through what Gary Watson calls the “aretaic perspective” (1996: 231). Plausibly, such judgments do not have to reveal anything about the agent’s character at all and does not imply that holding her responsible is appropriate. As an example of attributability responsibility, we can e.g. imagine performing some mundane action with no clear moral significance, which does not seem to be any reflection of the agent’s character, such as putting the left shoe on before the right shoe. In the normal case, where this is done consciously, it is clearly attributable to the agent as something she has done, and an outcome she has brought about. It seems clear that there is some - not just causal responsibility - sense in which it is appropriate to say, that this agent is responsible for this action and outcome.

When someone is attributability responsible, this at least implies that the individual in question is an appropriate candidate, in principle, for some further attribution of what Watson calls accountability. Accountability responsibility is the “holding” the individual responsible as in the practice of praising, blaming and punishing her.^{13 14} Accountability responsibility presupposed an ascription of attributability responsibility, but not vice versa. So when Strawson identifies moral responsibility with our praising and blaming practices, he is conflating at least these two notions, i.e. the practice of holding responsible, and ascribing an action or outcome to the individual. The implication Strawson makes, that the activation of the reactive attitudes constitute moral responsibility, seems to

¹³ An exception here may be that an outcome can be attributable to an agent, even if that agent can never meet the further conditions for moral responsibility, e.g. due to that agent not being sensitive to moral considerations.

¹⁴ There are arguably also other “faces” of responsibility than merely attributability and accountability. E.g. Scanlon (2008) argues for the existence of “Answerability” responsibility (Scanlon does not use the specific term himself, but others attribute it to him, e.g. Shoemaker 2011), which is a notion of responsibility related to attributability responsibility, but where even though it does not in itself justify sanctions, other parties are justified in demanding that the person answers for the particular action, trait or outcome, where the response (answer) from the relevant party may then justify accountability if it is unsatisfactory.

overlook the significance of this conceptually prior notion of attributability responsibility, and the fact that a different criterion, or different conditions, apply to attributability and accountability respectively. Incidentally then, the attributability/accountability distinction of responsibility seems closer to the ledger view of moral responsibility. However, instead of holding that there are conditions for moral responsibility and justifications for praise- and blameworthiness, we would say that there are certain conditions for attributability responsibility, and further conditions for accountability responsibility. In the rest of this project, I will assume the distinction between attributability responsibility and accountability responsibility is appropriate. This will simplify the task ahead. When testing whether the individual is morally responsible for policy outcomes, we can simply test whether she is attributability responsible first, and if she fails, ignore the question of whether she also fulfils the conditions for accountability responsibility. If we cannot attribute an outcome to the individual, she should not be held accountable for it.¹⁵

2.4 Conditions for moral responsibility, and the proposed strategy

In this section I will briefly specify the conditions which are usually considered necessary for attributing an outcome to the individual. This will be helpful in proposing a strategy for examining whether the individual is directly morally responsible for policy outcomes, in light of these conditions. I will explore these conditions in greater detail in the next two chapters. This section will progress in adherence with the attributability/accountability distinction from above. I will mainly focus on the former and only briefly touch on the conditions for accountability in the last part of this section. In the subsequent chapters, I will discuss what is further required for someone to be *accountability* responsible. I will argue that in order to attribute an outcome to the individual, the individual should have been able to exercise control over that outcome. As I will suggest,

¹⁵ Of course I have to note that if some argument can be presented, suggesting that accountability does not necessarily presuppose attributability, then this will potentially affect the conclusions I draw on the matter.

control requires firstly that the action was produced voluntarily, and secondly that the action actually brought about the outcome, or contributed to it. An action is produced voluntarily if it firstly was not physically forced by someone or something else, and secondly, if the individual fulfils a cognitive conditions. The cognitive conditions concerns whether she is aware that her actions contribute to a given outcome, and whether she is aware of the significance of that outcome. It also concerns when she may be excused for being ignorant of her contribution and the significance. An action brings about an outcome, or contributes to it, if the outcome depends on the individual's action. Put together, there are then three individually necessary and jointly sufficient conditions for direct attributability responsibility (just 'attributability responsibility' for the rest of this chapter) for an outcome, which can be stated roughly as such:

- a) The action that produced the outcome is unforced
- b) The individual fulfils a cognitive condition
- c) The individual brought about the outcome

I will specify the conditions in the rest of this chapter, however, again, I will not provide a wholehearted defence of the claim that these conditions are individually necessary and jointly sufficient in this chapter. I will simply show there is some plausibility to that claim, and then defend it to a greater extent in the next two chapters.

2.4.1 Attributability as control

Among many writers on moral responsibility, it is assumed that for an individual to be morally responsible for an outcome, she has to have exercised some form of control over that outcome.¹⁶ Put differently, and in line with the current exploration, we may say that in order for an outcome to be attributable to an

¹⁶ A consequentialist account of moral responsibility would be the exception here, since the question of whether the individual had any control is not necessarily relevant to whether praising and blaming the individual will lead to best consequences (Usually attributed to J. J. C. Smart 1961: 302). Blaming someone for an outcome, even though they lacked control over that outcome, may e.g. lead to that person acting in ways which will bring about best consequences, down the line.

agent, and thereby potentially make the individual morally responsible for that outcome, the agent has to at least have some control over whether or how that outcome comes about, or fails to come about.¹⁷ Thomas Nagel (Nagel, 1979) famously asserted that control was necessary for moral evaluation¹⁸, and according to Neil Levy, this is indeed the most commonly held view of moral responsibility today (2014: 109). Though there are exceptions to this¹⁹, it will in most ordinary cases seem plausible that the distinguishing aspect of whether we should attribute an outcome to the individual, is that she actually was able to bring about that outcome, or could have avoided bringing it about.

The idea that control is central to moral responsibility is compatible with both the Strawsonian view and the Ledger view above. It is thus not a further substantive view of moral responsibility, but rather a further specification of how to interpret a central aspect of these views, i.e. how to appropriately attribute an action or outcome to the agent in question. E.g., a reason why someone on the Strawsonian view may not be an appropriate target for reactive attitudes, could be construed as a lack of control e.g. because they are acting under strong abnormal psychological compulsion. Or, on the ledger view, an explanation for why the agent deserves a negative moral evaluation (a mark in her ledger) could

¹⁷ A precise definition of “control” is difficult to dig up in the philosophical literature. Daniel Dennett’s formulation is however at least suggestive, if vague: “A controls B if and only if the relation between A and B is such that A can drive B into whichever of B’s normal range of states A wants B to be in” (1984: 52). Of course, to “drive B”, and “normal range”, is quite vague. Dennett however suggests that the ‘normal range’ pertains to the foreknowledge the agent has, and that ‘drive’ pertains to that which we can actually initiate causally. So I have control over something, if my actions brought it about, and if I foresaw that my actions would lead to it.

¹⁸ But he nonetheless shows that almost everything we do and ascribe to the individual, is beyond our control, since it is susceptible to some kind of moral luck.

¹⁹ I am getting ahead of myself here. But I should mention, that aside from the more famous issues pertaining to moral luck, there are issues pertaining to culpable ignorance. Arguably (I will discuss it in chapter 4), we could not exercise control over an outcome which is the result of a literally thoughtless action, where we are not responsible for our thoughtlessness. However, some will argue that we can be morally responsible for such an outcome. If this is the case, then we can attribute an outcome an outcome to the individual which she we had not control over.

in part be because it is true that the person was actually able to produce the harmful outcome, i.e. was in control regarding whether that outcome obtained or not. The view that control is necessary for moral responsibility can historically be traced back to Aristotle who argues that a person is only an apt candidate for moral responsibility if the person acted voluntarily (Aristotle, 1998: Bk 3) “Voluntary” applies only to those things of which the person is the cause or origin, and in Aristotle's case it refers to an action and an outcome, or a trait (which I will ignore here). This plausibly ties into ‘control’, because, for the individual to be able to exercise control over some outcome O, it plausibly must at least have been the case that O was produced voluntarily. It however also requires more than that.

For an action to be voluntary, it has to fulfil two conditions, which are stated negatively. **Firstly**, the agent must not be compelled by an external physical cause (e.g. Broadie, 1991: 142, and Glover 1970: 1-21) to produce the relevant action or outcome. I will denote this '**upstream control**'. The reasons for this label is that this condition concerns whether the individual is in some way determined by antecedent external physical causes, i.e. something that is upstream from the individual. A clear example would be a gust of wind pushing someone, who as a consequence uncontrollably ends up harming someone else. Or it could be someone who *forces our hand* in the literal sense. In these cases the cause of her action is clearly external to the person in a way which inhibits the person's capacity for voluntary choice and her ability to exercise control over an outcome. Further, it also concerns manipulation, e.g. psychological or chemical manipulation, e.g. drugging or hypnotization where someone is manipulated to perform some action in a way which undermines her original will. We can also imagine a more illustrative case: Imagine e.g. person A walking down the street. All of a sudden a dog comes from out of nowhere, and runs out in front of her. Unable to react in time, person A trips over the dog, and falls into person B, who loses balance and falls to the ground, scraping her knees. Ignoring certain details, we would not attribute the outcome of person B scraping her knees to person A, as something person A did. There was no voluntary act, and person A did not clearly exercise control over the noted outcome. It was seemingly just an

unfortunate accidental occurrence. We could arguably also include something like the thesis of determinism here. I say ‘arguably’, because whether determinism falls under this condition, will presumably depend on one’s theoretical commitments with regards to the debate on free will and determinism. If free will is e.g. assumed to be compatible with determinism i.e. if the individual can perform the morally relevant sort of free action even if the universe is determined, then determinism does not necessarily undermine the individual’s control. However, as noted elsewhere, I will have very little to say on the topic of free will and determinism in this project.

Secondly, an action or an outcome has not been produced voluntarily if the person fails to fulfil certain cognitive requirements, i.e. a ‘**cognitive condition**’. This condition as noted concerns whether she had the intention to contribute to an outcome, in the sense that she was aware that she contributed to an outcome, and whether she was aware of the significance of that outcome. Further, it concerns when she is excused for lacking said awareness.²⁰ If she at the time of performance, was unaware that her action would produce said outcome, it is then appropriate to say the outcome was not brought about voluntarily. Imagine e.g. a person walking out the door as she does every day. However, on this day,

²⁰ A terminological note regarding ‘awareness’, ‘foreseeability’ and ‘intention’. I will assume ‘awareness’ and ‘foreseeability’ are synonymous, when it is awareness of a particular outcome obtaining. I will also understand ‘intention’ as meaning the same unless specified otherwise. ‘Intention’ is here understood as what Scanlon (2008: 10) calls the “broad sense”, which is just the opposite of *unintentional*. Again, as foreseeability. Another notion of intention, is intention in the “narrow sense”. This is what we usually mean by “aim”. One’s particular aim when performing some action, may play a role in affecting how we evaluate the severity of a person’s transgression when that person *also* foresees that her action will produce a certain harm. One’s aim is however not obviously a necessary condition for moral responsibility, assuming one can have an aim without foreseeability. If I drive on the sidewalk with the exclusive aim of getting to work more quickly, I can still be morally responsible for the destruction I cause, even though I never aimed for it. I can, because I foresaw that my action would bring about that outcome. It was not “unintentional” if I actually expected to cause this harm. On the other hand, we would perhaps feel justified in an even harsher reactive attitude, if we came to know that I actually *also* wanted to hurt people. So intention in the sense of foreseeability or awareness is necessary for moral responsibility, but aim is not clearly necessary, though it may affect our evaluation nonetheless.

unknown to her, her closing the door starts – in a subtle butterfly-effect kind of way – an extremely unlikely chain reaction which results in a car crash on the other side of town two weeks from now. This outcome was not something she can be said to have produced voluntarily, and by some stretch of the term, it was outside her control. She was unaware that the outcome would obtain, and she was unaware that her action would bring it about. Regarding the voting case, this pertains to whether the voter understands the significance of the policy outcome in question, and thus whether it is plausible that the voter could foresee the consequences of a particular candidate winning. It also concerns whether she has sufficient reasons to expect that her vote will ultimately make a contribution to the policy outcome, i.e. whether she was aware that her vote would contribute. If she was completely ignorant of a given outcome being a consequence of some political candidate winning, that outcome can presumably not be attributed to her (unless she ought to have been aware). Further, and more complicated, if she is unaware that her voting action brings about or contributes to the outcome, the outcome can presumably not be attributed to her either (unless she ought to have been aware). Again, it is not something she has brought about voluntarily.²¹

The **third** and last control condition which does not relate to Aristotle's discussion on voluntariness, is opposite to upstream control. It is the notion of downstream '**causal responsibility**' (originally stated as a necessary condition by Joel Feinberg (1968: 674)²²). Causal responsibility concerns whether we in certain circumstances

²¹ Recall again, 'control' concerns whether the individual was able to make a difference to whether an action was produced, or an outcome came about. To say that the cognitive condition pertains to whether the outcome came about is perhaps a stretch of the term. However, the idea is that if we did not expect, or had no good reasons to expect that our action would bring about an outcome, then we were unaware that the outcome depended on us. If we were unaware that it did, we were also unaware that we ought to act differently to avoid bringing about the outcome. Therefore, we were unable to avoid the outcome coming about. Again, it may seem like a stretch to say the individual could not exercise control over the outcome, but the substantive concern still applies. Because of our ignorance, we had no reason to change our behaviour, because we did not know it would lead to the relevant outcome.

²² I am sure the idea can be traced back to much farther. But I am not aware of anyone using that term.

actually did affect some outcome directly through our bodily movements, or in the case of an omission, whether we could physically have prevented the outcome (though causal responsibility and omissions are a contentious issue: see chapter 3). This also relates to what in common law is referred to as 'Actus Reus', or the external or objective element.²³ It concerns what the agent caused, i.e. it concerns whether an external causal relationship obtains between the agent (or her body) and the action or outcome. Regarding an 'action', it concerns whether the agent initiated the physical bodily movements. Regarding the 'outcome' it concerns whether the individual actually initiated the causal chain of events that produced the outcome. I will in this project ignore the former entirely. I will assume that the agent in question, i.e. the citizen, is physically the author of her own actions in the causal sense. If we return to the door-closing example from before, imagine instead that the person has a reasonable belief that closing the door will actually lead to this car crash on the other side of town, two weeks from now. We can e.g. imagine that closing the door at a specific time, is the "go ahead" signal to a hitman, who will cut the break lines of some specific person's car two weeks from now, in order to kill the person in the car crash. If she closes the door, but the hitman forgot their appointment, but the car still simply *happens* to crash at the same time, this outcome is not attributable to her. It was not within her control, because *that* causal chain broke down, even though the outcome obtained. Now, we may still consider her a very bad person for *trying*. Indeed, the *action* of closing the door with the reasonable expectation that it will lead to a car crash, may be sufficient for putting her in jail. But, it would be inappropriate to say that the outcome is attributable to her, and at least for that reason, it would be inappropriate to hold her morally responsible for *that* outcome. Causal responsibility pertains to whether the individual could actually make a difference to the voting outcome. If she could not, then she is not causally responsible, and the outcome cannot be attributed to her directly. A note on "directly": Causal responsibility is what makes direct moral responsibility "direct". It concerns whether the outcome depended on the action of the individual. In

²³ *Actus reus* is generally referred to as the 'external element' of a crime in law, which is usually interpreted as causal responsibility (G.R. Sullivan 1993). The other component is the *Mens rea*, which then is the guilty mind, i.e. the mental/cognitive component

chapter 6 and onward, I will examine indirect accounts. These accounts can be distinguished from the direct account, in that they try to ascribe moral responsibility to the individual for an outcome, in a way that does not rely on the outcome depending on her contribution.

For the sake of this project, upstream control is the least interesting. The individual citizen who votes at a general election is usually not physically forced to do so. Further, whether determinism is problematic is a matter not directly pertaining to this case of applied moral responsibility. The discussion of free will and determinism concerns globally whether *anyone* can be morally responsible at all, i.e. all cases of moral responsibility. It would be outside the scope of this project to engage with that debate, aside from noting its relevance. I will for the sake of this project simply assume that the individual fulfils the necessary freedom-relevant condition²⁴ for moral responsibility. If it e.g. turns out that moral responsibility is impossible because of determinism, this will constitute a problem for applications of the moral responsibility to all cases, not just with regards to policy outcomes. The ‘cognitive condition’ and the individual’s ‘causal responsibility’ however concerns whether a particular person, or group of individuals, are morally responsible in specific (or local) circumstances. Therefore these will need to be examined in greater detail, and applied to the voting case.

2.4.1.1 Preliminary strategy

In light of the above, an obvious preliminary strategy seems possible. Specifically, examine whether the individual fulfils the individually necessary and jointly sufficient conditions for attributability, in turn. If the individual citizens fails to fulfil any of the conditions, then she is not directly morally responsible. Recall, direct *morally* responsible for an outcome, requires both attributability responsibility and accountability responsibility for the outcome. If we cannot

²⁴ This is Fischer and Ravizza’s term (e.g. 1993: 8). It indicates that the individual has to be able to make – what should be considered – a free choice. Whether that is then the ability to have done otherwise (discussed in Frankfurt 1969), whether there has to be some hierarchical alignment with our first order and second order desires (Frankfurt 1971), guidance control (Fischer and Ravizza 1998), or libertarian free will is less important for this specific project.

attribute an outcome to the individual, she is not accountability responsible either. Therefore, the strategy can be restricted to focusing on the attributability responsibility first, and then postpone the question of the individual citizen's accountability responsibility till it is clear that she is actually in the running for such an ascription. The individual is attributability responsible for an outcome if she fulfils three individually necessary and jointly sufficient conditions, namely if she (here more specifically stated than in the beginning of the chapter:

- a) Has upstream control,
- b) Fulfils the cognitive condition,
- c) Is causally responsible for the outcome.

As noted, I will simply assume the individual has upstream control. So in the next two chapters, I will discuss the causal responsibility condition and the cognitive condition in greater detail. Then I will apply these conditions to the case of political participation – specifically the voting case – to see whether it is plausible that the individual citizen fulfils these conditions. If she fulfils both, and assuming she has upstream control, it is appropriate to attribute the outcome her. This opens her up for a further ascription of accountability responsibility and thus an ascription of direct moral responsibility for policy outcomes. Conversely, if she fails to fulfil either condition, then she cannot be morally responsible on a direct account.

Before moving on to the analysis in the next two chapters, I need to briefly touch on the *accountability* responsibility of the individual. So far I have merely stated the conditions for attributing an action or outcome to the individual. A further question is when the individual is accountable for that outcome, i.e. when we should *hold* her responsible for the outcome. In order for the individual to be accountable, she needs to fulfil the conditions for attributability responsibility. However beyond this, the outcome plausibly has to at least constitute a harm (we cannot be accountable for neutral outcomes). What 'harm' then means, will often be a matter of dispute, and depend on one's theoretical commitments. Therefore the conditions for when the individual is accountable, pertains to a larger discussion. Thankfully, we can ignore most of this complexity if we simply focus on

clear cut cases. A clear example could be a person shooting and killing an innocent person. This is clearly a harmful outcome. So, if that outcome is attributable to the shooter, among other things because the shooter was causally responsible for the outcome, and fulfils the cognitive condition (is aware of the consequences of her action), then it would seem she is accountable for the outcome, perhaps insofar as blaming her, or punishing her serves some further purpose (if we accept the ledger view)²⁵. In the current context, we could then focus on unambiguous harms, such as the outcomes of unjust wars, or racial discrimination. However, this is presumably still not sufficient. If the shooter was e.g. coerced into performing the action, this may at the very least excuse her. More generally, it seems the cost of defection should be expected to be reasonably low²⁶. What that means in specific cases, will of course be a matter of debate, and I will for that reason avoid fringe cases. In the shooter case, this could e.g. mean that someone threatened to kill *my* family if I did not kill someone else. The fact that such a threat exists, would presumably affect my accountability compared to a situation where there was no threat. Again, as noted, I will elaborate this in the chapter on the cognitive condition and throughout the whole project.

2.5 Summing up the chapter

In this chapter I firstly disambiguated the concept of responsibility, and centered my focus on direct retrospective moral responsibility. I.e., moral responsibility for an action or outcome that has happened, which the individual brought about or made a significant contribution to, directly. I then examined two important substantive interpretations of moral responsibility, namely the Strawsonian view and the Ledger view. I noted that with regards to their application, they are quite similar, and that they lend themselves to a natural distinction between

²⁵ Recall, on the Ledger view, our reactive attitudes and further moral sanctions such as punishment, depend on further justification. On the Strawsonian view on the other hand, our reactive attitudes are constitutive of moral responsibility, so here they cannot be distinguished from an ascription of moral responsibility.

²⁶ Or in cases of omissions, the cost of performing the relevant action that would have prevented the harmful outcome, should also have be reasonable low. I will elaborate discuss omissions in the next chapter.

‘attributability responsibility’ and ‘accountability responsibility’. This distinction allows for clarifying the conditions for either type of responsibility respectively. With regards to attributability responsibility, I argued that a natural way of understanding whether some outcome ought to be attributed to the individual, was in light of whether she had been able to exercise control over the relevant outcome. I then elaborated the particular conditions for control. They were upstream control, the cognitive condition, and causal responsibility. If the individual fulfils these conditions, then she has control over the relevant outcome. If she has control over it, it is attributable to her. If it is attributable to her, she may potentially be held accountable for it, if she is aware that the outcome is harmful (or ought to have been aware) and if the cost of defection is sufficiently low. If she is both attributability responsible and accountability responsible, then she is fully fledged morally responsible for the relevant outcome. However, as I will show in the next two chapters, and sum up in chapter 5, in the case of voters and policy outcomes, the outcome is not clearly attributable to the voter. She is not directly morally responsible for policy outcomes. In the next two chapters I will examine the causal responsibility and cognitive condition in further detail (while ignoring upstream control).

3 Causal condition

3.1 Introduction

In this chapter I will examine whether the individual is causally responsible for policy outcomes by way of voting. First (3.2), I document that causal responsibility plausibly is a necessary condition for directly attributing an outcome to the agent, and thus a necessary condition for direct moral responsibility for outcomes. In doing this, I highlight that ‘omissions’ may constitute a problem for the strategy introduced at the end of chapter 2, and propose a possible ways to respond to this. Then (3.3) I examine whether the individual citizen fulfils the condition of causal responsibility for policy outcomes, when she participates in the political process by way of voting. In doing this, I highlight salient ways the individual’s causal responsibility can be undermined. I single out the problem of

overdetermination as the most significant challenge to causal responsibility in the voting case. In section 3.4. I focus on the problem of overdetermination, which is a problem due to the particular threshold structure of the voting case. I propose a solution to the problem, but then show that this solution leads to counterintuitive results. In 3.5 I sum up this chapter, and conclude that we lack clear grounds for holding the individual voter causally responsible, and thus for attributing the voting outcome to the citizen. If we accept this, this undermines direct moral responsibility for policy outcomes.

3.2 Causal responsibility as a necessary condition for direct attributability responsibility, and the problem of omissions

In this section I will examine whether causal responsibility for an outcome is plausibly a necessary condition for attributing an outcome to the individual directly.²⁷ ²⁸ I will start out by arguing that there seems to be an implicit acceptance of something like this condition in many of our judgments about moral responsibility, which seems to indicate that causal responsibility in part grounds moral responsibility. I will then cast doubt on the necessity of this condition, by highlighting the problem of omissions. Lastly, I propose how to respond to this problem.

²⁷ Recall, under the current framework, moral responsibility should be understood as composed of two necessary components, namely attributability and accountability. When an agent is attributability responsible for an outcome, the outcome is tied to the agent as something she e.g. voluntarily and knowingly did. If she is attributability responsible, this opens her up for an ascription of accountability responsibility. If she is accountability responsible, it is appropriate to *hold* her responsible for what she did, e.g. blame and punish her. Obviously then, if she is not attributability responsible, she is not morally responsible.

²⁸ Recall, *direct* moral responsibility involves that the outcome is also attributed to her directly, i.e. in light of her direct contribution to the outcome. Whenever I write “moral responsibility”, and “attributability” in this chapter, it should be assumed that I am referring to the direct notion of the concepts, unless I state otherwise.

When an individual is causally responsible for an outcome, this means straightforwardly that the individual brought about that outcome. This is usually interpreted as a certain relationship obtaining between a person (or a thing) and the outcome, in a way where it is appropriate to say that the person caused the outcome. How this relationship between the individual and the outcome is then appropriately spelled out more specifically, is a matter of debate which I will examine throughout this chapter. Concerning why causal responsibility is then often assumed to be a necessary condition for moral responsibility for an outcome, seems perhaps just as straightforward. This goes hand in hand with the concept of attributability responsibility highlighted earlier, where an action or outcome has to be tied to the agent in some way, in order for the individual to be morally responsible for that action or outcome. If we accept that moral responsibility presupposes attributability, causal responsibility can then be interpreted as partly constitutive of this *tying* the outcome to the agent as e.g. Joel Feinberg suggests (1970: 130)²⁹. Outcomes are things in the actual world, a subset of which are moral outcomes³⁰, and if the agent is morally responsible for something in the world, we could then see causal responsibility as part of what ties the agent to those outcomes, making her attributability responsible for them, and potentially morally responsible. Carolina Sartorio also mirrors this sentiment, and proposes that moral responsibility entails causal responsibility (2007: 750) and in part grounds moral responsibility, because causal responsibility tells us something important about the individual's relationship to the outcome. Even if causal responsibility is necessary for attributability responsibility, and moral responsibility, most would agree that some strong causal link between a person's behaviour, and a harm, is insufficient for an ascription of moral responsibility (e.g. as documented by Szigeti, A 2014 and Sartorio, C 2007) since such a relationship can e.g. be entirely accidental. To give an obvious example of this, say e.g. that future scientists establish with absolute certainty, that my writing this sentence causes the near extinction of the human race 200 years from now.³¹ It is doubtful

²⁹ As I will elaborate later on, it is not the only possible interpretation of this link.

³⁰ The "moral" part is not necessarily in the world in a sense that implies moral realism.

³¹ Joel Feinberg and H.L.A. Hart have both given an account of the centrality of this condition both for moral responsibility, and its legalistic application and rationale (Feinberg, J. 1970 and Hart, H.L.A.

that any surviving future reasonable person with this information in hand, would consider me morally responsible for this outcome based *solely* on this fact about the causal relationship between me and that outcome. More is needed.³² If I am merely causally responsible, that outcome is not attributable to me as an agent, but merely attributable to me as a thing (e.g. my fingers punching keys). Because it is insufficient for attributability responsibility, it is thus also insufficient for direct moral responsibility.

In order to shed light on whether causal responsibility is a necessary condition for moral responsibility, I will need to examine the particular notion of causation which is at the heart of that condition. The most common understanding of “cause” with regards to moral responsibility today, is the counterfactual analysis of causation³³. On the counterfactual analysis, in its simplest form, one event causes another event, if the latter event depended on the former (Lewis, 1973)³⁴. To give an example of this analysis in action, we might imagine the following: the lightning struck the thatch roof and the house burned down. If lightning had not

1968). For a more recent discussion of this conditions see Sartorio, C. (2007). Some account of causal responsibility can be traced all the way back to Aristotle, in his deliberations on voluntary and involuntary actions. (Echeñique, J 2012).

³² As I will argue in the next chapter, part of what is needed is some degree of awareness on my part that my actions will make a causal contribution to such a harm.

³³ originally sketched by Hume (2003: 122), but thoroughly developed by Lewis, D. 1973. Further, its popularity is documented by e.g. Bennet, J. 1987, Goldman, A. 1999 and Sartorio 2004. An alternative theory of causation which is the regularity analysis. However, the regularity analysis has fallen out of fashion lately, and is certainly less pervasive in writings on moral responsibility. Further, there are various substantive criticisms against (e.g. it fails to distinguish between causes and mere correlates (Kutz, C. 2007)), which I will not go into here. Further, since this analysis shares the same drawbacks as the counterfactual model of causation with regards to the voting case, I will not give it any significant attention due to lack of scope. But, if someone were able to give a positive account of the causal responsibility of the individual voter, based on this account, this could of course potentially affect the verdict in this case.

³⁴ This is stated in its simplest form, which is not without problems. There are other ways of stating this, e.g. probabilistically, where the *probability* of the consequent event obtaining would vary if the antecedent event obtained (Lewis, 1986: 120-1). I will however focus on the simple account here for purposes of clarity.

struck the thatch roof, the house would not have burned down. Or to take a clearly morally significant example: Paul shoots Susan, and Susan dies from the injury. Had Paul not shot Susan, Susan would not have died from the injury. In these examples, the lightning strike caused the fire, and Paul caused the death of Susan, because these outcomes were dependent on the relevant actions. The counterfactual analysis is by no means unproblematic, but I will not attempt to engage with any of the problems in any great detail.³⁵ However, at least two features of this analysis make it a relevant candidate for an account of causal responsibility. One, it does seem to capture our ordinary way of thinking about causes in many instances, both with regards to agents, and with regards to outcomes in the physical world. The reason is that it attempts to pinpoint the exact cause of something in a clear way as indicated by the previous examples. Two, it seems, as noted, that the counterfactual model of causation, or at least something close to it, is implicitly accepted by many writers on moral responsibility when accounting for moral responsibility for outcomes, or when highlighting problems with doing so.

If something like the counterfactual analysis explains what we mean by 'causal' in causal responsibility, causal responsibility as a necessary condition and a grounding claim may be problematic right off the bat. The problem relates to omissions. The problem is that it seems clear that we can attribute outcomes to individuals, which come about in light of omissions. However, some argue that omissions are not clearly causes. If they are not causes, but we can attribute outcomes of omissions to agents, then causal responsibility cannot be a necessary condition for attributability. If causal responsibility is not a necessary condition for attributability, we need to revise the assumption from chapter 2, that causal responsibility, upstream control and the cognitive condition are individually necessary, *and* jointly sufficient for attributing an outcome to the individual.³⁶ If

³⁵ It is simply outside the scope of this project to defend the counterfactual analysis and it would get in the way of actually applying causal responsibility to the case. See e.g. Beebe, Hitchcock, & Menzies, 2009 for a comprehensive overview of the discussion.

³⁶ Let me sum up the assumption from chapter 2. The individual's direct moral responsibility for an outcome, presupposes that an outcome can be attributed to the individual (attributability

they are not individually necessary and jointly sufficient, we have to revise the overall applied strategy for this project in some way. So far the strategy was to simply assume that the individual had upstream control³⁷, and then examine in turn whether she fulfilled the other two conditions. If she did, then we could attribute the outcome to her, opening her up to an ascription of accountability responsibility, and thus moral responsibility. If she failed on just one of the two conditions, then we could conclude that she was not morally responsible. However, if omissions are not causes, but we can be directly morally responsible for the outcomes in light of omissions, this means that causal responsibility is at best a part of a necessary set of disjunctive conditions. If that is the case, we cannot exclude her moral responsibility for policy outcomes by simply ruling out her causal responsibility. Therefore, we need to examine omissions in more detail, to see whether they indeed constitute a problem for the claim that causal responsibility is a necessary condition for direct attributability responsibility.

Firstly, it does indeed seem intuitively clear that we can attribute an outcome to the individual in light of an omission. Example: I have promised to watch my neighbour's cat for a month while she's away on holiday. During that month, I am too lazy to feed the cat, which then is dead from starvation when my neighbour returns from holiday. It seems clear that I can be morally responsible for the *outcome* of that omission, i.e. we can attribute the horrible condition the cat is in to me, even though it is the result of an omission, and not a positive action on my part. However, the idea that omissions can be causes is contentious (The contention e.g. documented in Sartorio 2004, and Beebe, Hitchcock, & Menzies, (2009)). Some will claim in the cat example from before, that there is no obvious action or outcome which actually occurs and originates in me, which has the causal power of producing the outcome of the cat's condition, and which – if I had

responsibility). As suggested in chapter 2, causal responsibility, upstream control, and the cognitive condition are individually necessary, and jointly sufficient for attributing moral responsibility to the individual.

³⁷Recall, to have upstream control means that one is not forced by antecedent circumstances (e.g. physically pushed) and fulfilled the freedom-relevant condition pertaining to the discussion of free will and determinism.

not done it – the outcome would not have obtained. This rests on the understanding that only events can be causes (have causal power), where events are defined as “positive occurrences with definite spatiotemporal locations” (Sartorio 2007: 752). An absence then does not conform to that definition (Dowe 2001). If we accept that absences are not causes, we may still be able to give some counterfactual explanation for how the cat died, or how a broader causal pattern of events obtains which bring about the outcome of the cat dying, and where the cat would not have died if that pattern of events had not obtained.³⁸ What we lack however, are clear grounds for claiming that *my omission* was the cause of that outcome.

In spite of omissions seemingly lacking causal power, many will still share the intuition that the outcome of the cat dying can be attributed to me. However, if omissions are not causes, and if causal responsibility is a necessary condition for attributability, then the cat example, and other examples like it, seem to rule out attributability responsibility in omission cases. Indeed, since our intuitions in omission cases strongly suggest that we can attribute the outcome to the individual. This has led some writers to conclude that causal responsibility is not a necessary condition for attributability at all (See e.g. Driver p423 in Sinnott-Armstrong 2008). This conclusion, though tempting, comes in conflict with the intuition that there is a significant *moral* difference between attempting to bring about O, and actually bringing about O. Imagine three roughly sketched scenarios involving me and an innocent target:

- 1) I shoot my gun at someone, and miss the target.
- 2) I shoot my gun at someone and hit the target, killing the target.
- 3) I shoot my gun at someone, miss the target, but they die of an unrelated heart attack at the same moment the bullet passes by the target.

³⁸ Lewis e.g. holds this view, and argues that though we can explain the relevant omission-related outcomes in terms of certain patterns of positive events, it is inappropriate to claim that any non-event *caused* the particular outcome (Lewis 1986: 188).

Assume the only difference is the causal component. Presumably our intuitions will vary across cases. In case 1) I am clearly not morally responsible for the outcome of the death of the target, because the target is simply not dead. In 2) I *am* clearly morally responsible for the death of the target, and the most obvious explanation is that this is simply because I caused her death. In case 3) I am not morally responsible for the death of the target, even though she is dead. The distinguishing feature with regards to when I am morally responsible for the *outcome* of the target dying, seems intuitively to be that I actually bring about the outcome i.e. my causal responsibility. We may find it appropriate to throw me in jail in all three cases, because I am a dangerous person, who tries to do bad things. But only in one case does it seem appropriate to say that I am morally responsible for the *outcome*. Only in one case is the *outcome* attributable to me. Causal responsibility seems, at least in some cases, to *ground* moral responsibility. This suggests that actually causing the outcome makes a clear moral difference to when an outcome can be attributed to the agent. If we accept this, it seems appropriate to examine whether omissions can be incorporated into the concept of causal responsibility.

One proposal of how to incorporate omissions into causal responsibility is to include what Phil Dowe calls 'Quasi-causation' as a sufficient condition for causal responsibility. To be the quasi-cause of something, instead of a fully-fledged cause, then suffices for causal responsibility for it. We should understand this as a supplement to the counterfactual model of causation, and not an alternative model as such. This supplement then attempts to explain the intuitions we have in omission-type cases that the individual produces the outcome. Quasi-causation obtains, where the mere possibility of causation is a kind of causation. Phil Dowe claims that in many cases this grounds our intuitions of causation with regards to a non-event. With regards to attributability, quasi-causation and causation are then disjunctive necessary conditions for causal responsibility. I.e., an agent is then attributability responsible for O, if the agent is causally or quasi-causally responsible for O. (Sartorio, C, 2007: 53). On this account, if it is true that had I fed the cat, it would still have been alive, then I am quasi-causally responsible for its death when I do not feed it. The problem with this view is that quasi-causation

and thus causal responsibility overgeneralises as it is simply too easy to come by, which thus seems to trivialise causal responsibility it as a grounding claim. The queen of England is quasi-causally responsible for the 9/11 attacks, because if she had called the CIA a week before the incident and explained to them in detail about the impending attack, presumably (or let us just assume), the incident would have been averted. The cat not spontaneously developing the ability to shop for food at the local pet store is causally responsible for its starvation. Pol Pot not coming back to life, breaking down the door, and feeding the cat is also casually responsible. In this case, causal responsibility is too easy to come by, and thus loses its significance as a grounding claim for moral responsibility. The inclusion of quasi-causation under causal responsibility hardly salvages causal responsibility as a necessary condition for moral responsibility³⁹. It trivialises the condition because it is too easy to fulfil. The assumptions that causal responsibility in part allows us to sort through which outcomes should be attributed to the agent loses its appeal when almost anything can be a causally responsible for an outcome.

The problems with accounting for causal responsibility in omission type cases, has motivated an important response. Instead of focusing on metaphysical causation as a basis for causal responsibility, the strategy should instead be to try to link the omission to a normative aspect of the agent in question, specifically our intuitions of moral responsibility, as a way of tracing out the relevant causes from the irrelevant ones (Thomson, J. 2003 and Smiley, M. 1992). The reason the queen of England is not causally responsible for the 9/11 attacks, is because appropriate ascriptions of moral responsibility trace out the normatively relevant notion of causing something. So if I promise to water your expensive plants while you are away and the plant dies, this outcome is clearly attributable to me, even though it is an omission. It is attributable to me because that particular omission should in this case actually be understood as a cause. The reasons it should be understood as a cause is because our intuitions of moral responsibility trace out the moral significance of this particular omission as something *I* brought about. As Marion

³⁹ A version of The Queen of England case is not mine, and is used by quite a few different writers (Beebe, Hitchcock, & Menzies, 2009).

Smiley writes, when determining causal responsibility, “though we may think of ourselves as discovering the objective causes of harm, we in fact import into our causal analysis an assortment of more purely conventional attitudes about whose interests count in society” (Smiley 1992: 255). Expanding somewhat on this view, the question of what an appropriate causal relationship is, then takes a back seat to other aspects of our analysis of moral responsibility, at least in some cases. Not necessarily to the extent that it lets us attribute causal relationships where there clearly are none, but the particular relevant notion of a causal relationship necessary for moral responsibility, may to an extent be a matter of convention, rather than metaphysical causation. And – this is the claim – since our normative intuitions in a given case suggest that some omissions are important causes, this lends credence to such a claim based on this type of analysis. This then removes the fear of potential omissions overgeneralizing, because not all omissions count, e.g. not the Queen’s omission. The Queen is not morally responsible, thus she is not causally responsible.

Though this sounds promising, it causes a potential problem for the strategy of this project, and thus with answering the overall question. Causal responsibility may still be a necessary condition for attributability responsibility. However, if we have to rely on intuitions of moral responsibility to determine what it is the individual is causally responsible *for*, it becomes difficult to ascertain whether the individual is morally responsible in the first place, in cases where we lack clear intuitions of moral responsibility going in, as in the voting case. Recall, the strategy was to examine whether the policy outcome could be attributed to the citizen via voting. If it could, this would then open her up for an ascription of moral responsibility.⁴⁰ Specifically, in order to test whether she was attributability responsible, the strategy was to examine whether she fulfilled the individually necessary and jointly sufficient conditions for attributability. But, if one of those individually necessary conditions can only be ascertained if we already have a clear sense that she is morally responsible, that undermines this strategy (because we do not have a clear sense). In cases where we have unclear intuitions going in,

⁴⁰ Recall, moral responsibility plausibly had two parts: attributability- and accountability responsibility.

such as in the case of individual moral responsibility for policy outcomes, we need a strategy for deriving a verdict. It would be helpful if we could focus on the question of causal responsibility, and *then* use any conclusions regarding the citizen's causal responsibility to shed light on the overall question of whether the citizen is then morally responsible. However, this is not possible in the current case, because we have no clear means of ruling out her causal responsibility. There could always be some omission which could be traced out if we had clear intuitions of moral responsibility.

A possible way to save the strategy is to simply concede that our moral intuitions are instructive in sorting through all the possible omissions which are possible causes of the outcome. However, this does not rule out that we can then still hold that in cases of positive actions (or cases of '*positive agency*' as Fischer and Ravizza calls it, 1998: CH5), we can still determine the cause without clear moral intuitions going in, in light of the counterfactual analysis of causation. In cases of positive actions, there is usually not an extreme range of possible relevant causes to choose from, so we do not need our intuitions of moral responsibility to sort through them all. In light of this, we can then for the sake of this project, focus on cases where the individual performs an actual positive action, and see whether she is then causally responsible for policy outcomes via voting, in accordance with the strategy proposed in the previous chapter. It is however still important to highlight that this constitutes a limitation to the strategy. Recall again: the strategy was to examine whether an outcome was attributable to the individual. If it was, then she was open for an ascription of accountability. If she was then also accountable, she would then be morally responsible for the outcome. So the first step was to examine whether the outcome was attributable to her. As suggested in chapter 2, it would be attributable to her if she fulfilled three individually necessary and jointly sufficient conditions. Specifically, the freedom-relevant condition (assumed fulfilled), the causal responsibility condition, and the cognitive condition. Therefore, if we can show she is not causally responsible, then we can rule out her moral responsibility, because the outcome is then not attributable to her. However, if we are unable to evaluate whether she is causally responsible for the outcome in light of an omission, since we lack the moral intuitions to sort

through all the possible omissions, then we cannot rule out her causal responsibility, even if we can rule out that she is causally responsible in light of a positive action. Any negative conclusion about her direct moral responsibility will then contain the caveat that it only pertains to positive actions, not omissions.

Summing up this section: causal responsibility is plausibly a necessary condition for moral responsibility for outcomes and causal responsibility seems to ground moral responsibility. Further, causal responsibility can presumably be understood in accordance with the counterfactual analysis of causation. A contention in relation to this was that omissions, or rather absences, are not clearly causes. It however still seems clear that we can be morally responsible for outcomes in light of omissions. If this is true, and if we accept that we can be morally responsible for outcomes of omissions, then causal responsibility is not a necessary condition for moral responsibility. The solution to this was to accept that omissions were quasi-causes, and then accept that causation and quasi-causation are necessary for causal responsibility disjunctively. The acceptance of omissions as a *type* of cause produced a problem for the strategy introduced in the previous chapter. This strategy was to test whether the individually necessary and jointly sufficient conditions were satisfied in the voting case. However, because quasi-causation overgeneralized, we needed some way of sorting out the relevant causes. The suggestion adopted from Thomson and Smiley, was to employ our normative intuitions to trace out the relevant causes for testing. However, since we do not have any clear intuitions of moral responsibility in the current case, we cannot trace out any relevant omissions, even if there would be some if we had clear moral intuitions. So simply showing that the individual does not produce the outcome through her voting action, does not *in principle* rule out her causal responsibility in light of this analysis. The analysis can however still be applied to produce a positive verdict. If the individual voter is causally responsible for policy outcomes, then the condition is satisfied, potentially grounding an ascription of moral responsibility. Further, it can also be applied to examples of positive action, and we can at least potentially rule out that she is causally responsible in those cases.

3.3 How causal responsibility can be undermined

Assuming causal responsibility in positive action cases is a necessary condition for attributability responsibility, there are two obvious ways causal responsibility can be undermined. The first pertains to whether the outcome even obtains. If e.g. the policy outcome never comes about, the citizen can clearly not be causally responsible for it. There are many reasons why an outcome does not come about. E.g. the relevant policy (or relevant candidate) may simply have failed to acquire a sufficient number of votes. Or perhaps no political candidate wanted to make it part of her base. Or, as is the case in proportional representation systems, the relevant party or politician may try to get elected with the promise of implementing a given policy. However, due to these systems tending to produce minority governments, she, or her party, usually has to broker alliances with other parties. In that case, it is not given that she will be able to find agreement on that particular policy. The second obstacle is more interesting. This obstacle constitutes a challenge *even* if the outcome obtains. It pertains to whether the individual can be said to have made a contribution at all, in voting-style “threshold” cases (more on thresholds later). This last obstacle concerns the problem of ‘overdetermination’. Overdetermination occurs when more than a sufficient number of causes are in play. When they are, it is not obvious that the outcome depended on any of the causes, taken individually. If the outcome did not depend on any of the individual causes, this seems to lead to the conclusion that none of them causes the outcome. Applied to the voting case – if no individual voter causes the outcome, then no individual voter is causally responsible for it. If causal responsibility is a necessary condition for direct attributability responsibility, then this leads us to the conclusion that the voter cannot be directly morally responsible for the voting outcome, and thus not the policy outcome of which the specific voting outcome was a precondition. The second problem is more interesting, and requires a more expedient solution, since it concerns whether we can be causally responsible in voting cases *in principle*. I will focus entirely on the problem of overdetermination, and whether it undermines causal responsibility. The previous obstacles narrow the field of what particular outcomes a voter can be causally responsible for, if she can be causally

responsible in principle, but they do not undermine her causal responsibility in principle.

I will firstly elaborate what is meant by overdetermination, and why overdetermination is a problem for an ascription of individual causal responsibility. Then I will show how overdetermination is a problem in the voting case specifically. Lastly, I will show that when the voting case is interpreted as a specific type of overdetermination case, namely as a pre-emption case, this allows for an ascription of causal responsibility. Lastly, I will conclude that contrary to first appearances, this solution is not at all promising. We therefore seem to have to conclude that the voter is not causally responsible for policy outcomes, due to the problem of overdetermination, at least, if we ignore the possibility of some omission being a relevant quasi-cause.

3.3.1 Introducing symmetrical and asymmetrical overdetermination

Overdetermination comes in two varieties, 'symmetrical' and 'asymmetrical' and whether we should interpret the voting case as one or the other, may affect whether we can ascribe causal responsibility to the individual voters. Symmetrical overdetermination occurs when a particular event is caused by two or more causes, and where more than enough causes are active and thus sufficient for bringing about the event, and where the event had not obtained if none of the causes had been in effect. An example of this is the classic firing squad case of multiple (e.g. 10) shooters all hitting at the same instance, in the same place, but where a smaller subset (e.g. three) of them hitting would have been sufficient to kill the victim. Because more than the required number of shooters hit, the outcome is symmetrically overdetermined. It did not depend on any single individual (even though it depended on a group). Asymmetrical overdetermination on the other hand occurs when more than enough causes are active, and thus sufficient for bringing about the event, but where there is a discrepancy between the causes, which allows us to single out certain causes as more or less significant to bringing about the outcome. This would e.g. be the case

when some of the shooters have bigger guns, or where there is a temporal discrepancy between the individual shooters (Lewis, 2000: 182). In the following, I will examine whether the voting case should be interpreted as one or the other, and why this is relevant.

On the counterfactual analysis of causation, the individual does not obviously cause the outcome in cases of symmetrical overdetermination. The reason is, had she not fired her rifle, the victim would still have been killed. This does of course not mean that a conjunction of causes did not bring about the outcome (Lewis 1986: 199 and 2000: 182). If a sufficient number of shooters had not fired, the outcome would not have come about, so we can clearly conclude that the outcome depended on a number of shooters firing their weapon at the target. But since we are trying to account for the causal responsibility of the individuals *as individuals*, this in itself does not straightforwardly mean that the *individual* is directly causally responsible, just because *the group is*, or a subset of the group of individuals are. The problem is that it seems we are forced to say that the individual participant is not causally responsible for the outcome *as an individual*. The outcome did not depend on her contribution. The exact same outcome would have obtained whether she participated or not. If the voting case is then an example of symmetrical overdetermination, that is a problem. It is a problem because if causal responsibility is a necessary condition for direct moral responsibility in cases of positive actions, then the individual is simply not causally responsible. If she is not causally responsible, we cannot attribute the outcome to the individual. If we cannot, then she is not morally responsible in light of the strategy specified in chapter 2.

As noted, in the case of asymmetrical overdetermination, contrary to symmetrical overdetermination (Lewis, D. 1986: 171, Moore, M. 2009: 86), something separates the causes. I will focus on one such notion of separation, namely the temporal. This is also known as 'pre-emption', precisely because one cause temporally pre-empts another. Here we have two or more potential causes, where one actually causes the effect, but if it had not, some other cause would have (Lewis 2000: 182). A case of asymmetrical overdetermination could then be

where there is a small temporal discrepancy between the shooters in the firing squad, but where one shooter's bullet reaches the victim, and kills him, prior to the other bullets arrival, but where insofar as she had missed, someone else would have hit the target and dealt the killing shot.⁴¹ In the pre-emption case, it seems at first glance also to be the case that the individual's contribution does not matter. The reason is, the individual, even though her bullet reached the victim first, can still claim that the outcome would have obtained even if she had not fired her rifle. It did not depend on her shot. However, though this would have been the case under symmetrical overdetermination, it is not necessarily true in the pre-emption case. According to e.g. David Lewis, there are clearer grounds for the claim that at least some of the participants are individually causally responsible on the pre-emption case, than in the case of symmetrical overdetermination.

The reason pre-emption cases are less problematic cases of overdetermination, is that we are permitted to understand causation in terms of a whole causal chain, not just a single cause and a single outcome. In the pre-emption case, though it is true that if the first shooter had not killed the victim, the second shooter would have, it is – or so the claim goes - not true that the same causal chain would have obtained. It would have been a different causal chain altogether that would have obtained. Because of this, it would have been a different *event*, or rather, a different *compound* event altogether. David Lewis refers to this permissible variability in the description between events as “fragility” (Lewis, D. 2000). If we accept this, then *some* shooters can still be the cause of the event, specifically the first x number of shooters sufficient for bringing about the outcome. If it took three shooters, but 10 actually shot, then we may be allowed to say that the first three are causally responsible. Specifically, imagine I am a shooter in the firing squad. Assume that it took three hits to kill the victim, and assume that I was

⁴¹ Overdetermination may be asymmetrical in other ways than in the temporal aspect. E.g. some causes may be stronger than other causes. E.g. in a tug-o-war, the number of people pulling on one side, and the force they transmit to the rope may be more than enough for bringing about the outcome of their side winning, but presumably some of the rope pullers are also going to be stronger than others, meaning that they will have a different effect on the rope.

shooter number three. Assume also that there is a temporal discrepancy between the shooters. In this case, I (and the other first two shooters) am the cause of a specific causal chain obtaining. This is the case because *that* chain had not obtained if I had not shot. Had I not shot my gun at the victim, event A, or causal chain A, which among other things includes me hitting the target, and the victim dying, would not have obtained. Instead another event B, or causal chain B would, instead have obtained. That event would have contained a different configuration of shooters. I am thus causally responsible for event/causal chain A, because that event depended on me shooting.⁴² If the voting case is a case of asymmetrical overdetermination, then we may have grounds for holding the voter causally responsible.

Symmetrical overdetermination on the other hand does not clearly allow for such a liberal description of the event. Due to the symmetrical nature of the case, there is no relevant discrepancy (e.g. temporal) between the causes which allows for an alternative counterfactual construal of the whole causal chain as a separate event, which includes the outcome. It is exactly defined as a case where all the causes are symmetrical. If construed as a whole chain, that is it. If that chain does not obtain, then the whole causal chain fails to obtain. So the individual's contribution has to be evaluated within *the* particular chain of events that obtains, and goes through the individual. And, in that particular chain of events, the individual's contribution is overdetermined. Thus the outcome did not depend on the contribution of the individual, and she is not causally responsible.

⁴² A point to note here, is that not every asymmetry will count as relevant, and events can be individuated too fine-grained, as e.g. Moore, M.(2009) points out, when he asserts that it generates "enormous promiscuity" (ibid: 413). E.g. we can imagine a scenario where 10 people act in a way which seemingly overdetermines some outcome symmetrically, e.g. it only took five to produce the outcome. We then come to know that one of the individuals as the only individual wore a hat. So seemingly, if he had not participated, it would have been a *completely* different event, because it would then not have been a compound event which included one hatted participant. So the hatted person actually caused the whole event, right? Clearly not. The hat is intuitively inconsequential to the description of the event.

3.3.2 Voting – a threshold case

The question is now whether we should understand the voting case as an overdetermination case at all. I will argue that we should. The question then becomes whether it is a case of symmetrical or asymmetrical overdetermination (pre-emption). If it is a case of symmetrical overdetermination, then we seem forced to conclude that the individual is not causally responsible, since the outcome does not depend on her contribution. If it is instead a case of asymmetrical overdetermination, the individual may potentially be causally responsible.

In order to answer whether it is an overdetermination case, I first need to elaborate the voting case itself. In order to do this, it will be helpful to distinguish between two different types of cases. Voting is a case of collective action in the plain sense that it involves more than one person. Collective action cases can roughly be sorted in to two categories (Tuck 2008: 30). Firstly, there are variable outcome cases. These are cases where the outcome is contingent on the contributions of many, but where every individual contribution has some – if only minute – effect on the overall outcome. So if we take global warming as an example, we can say that the amount of CO₂ the individual contributes to the atmosphere, will theoretically – everything equal – have some very minute effect on the global temperature. Secondly, there are threshold cases. A threshold case is one relating to the proverb of the straw that breaks the camel's back. In this sense, there is a certain threshold which has to be filled until some specific outcome obtains. Anything short of that, or beyond that threshold, does not contribute to the outcome. If the threshold is not met, the outcome of the back breaking does not obtain⁴³. If it *is* met, the camel's back breaks, but it does not – let us assume - break to a larger extent, or once more (assuming it is a binary relation). So when we want to tie the agent's contribution to some outcome in collective action cases, it may in some cases relate to the first category (variable

⁴³ The example is not entirely apt, because in the unfortunate camel's case, it *does* actually also contribute to the strain on its back. In a regular threshold case in the sense I am using the term, this is not the case.

outcome case) in the sense that it is a part (e.g. a small part) of the whole outcome. In this case it is then at least this small part the agent is causally responsible for (and potentially morally responsible for). If it is the second type, i.e. a threshold case, it is a matter of debate what the individual's contribution is. I.e., it is a matter of debate what the individual is causally responsible for. Overdetermination does not occur in the variable outcome case, but it can occur in the threshold case. It occurs if more than enough contribute to filling the threshold. Questions of who is causally responsible for what, seem immediately easier to settle in variable outcome cases, because we can see the direct effect of the agents contribution to the relevant outcome. Even if the contribution is a minute contribution, it is at least something she causes.⁴⁴

Because the voting case is a threshold case, it is an apt candidate for being an overdetermination case, precisely because in all but the most unlikely of scenarios, the number of votes that are cast, exceeds the threshold. There is one extremely unusual scenario in which the threshold is met, where it is not a case of overdetermination. This case is one where the threshold is met *exactly*. So assume 40 people vote for candidate A, and 41 people vote for candidate B. In this case the winning threshold is met. But, only by one vote. The interesting thing to note is that the outcome is here dependent on all of the voters. Taken individually, holding the acts of the other voters constant, the outcome depends on each voter actually voting. The case is thus not an overdetermination case at all. The outcome depends on each individual voter contributing, individually. The threshold requirement is exactly met. Counterfactually, if only one of the voters – Paul – had failed to vote, the voting outcome would have been a draw, and B would not have won. Therefore, Paul, taken in isolation, is the cause of the

⁴⁴ As Jonathan Glover (1975) argues, the variable outcome case can seemingly be construed as a threshold case of sorts. Specifically, it can be construed as a threshold case in the sense, that there is a threshold to be met before a discernible difference in the outcome obtains. If e.g. a cubic inch of CO₂ is expelled, the impact to this on global warming will presumably be so small that no person or animal alive now or in the future, will register the harmful effect of such a minute increase in global warming. So in that sense we may talk about a threshold even in variable outcome cases, i.e. a threshold of compound significance.

outcome on the counterfactual analysis of causation. But not just Paul of course. This is also true for every single voter who supported B, taken in isolation, holding everything equal. Everyone is a “pivotal voter”. The odd thing is of course, that had just one more voter, Tim, decided to vote for B that day instead of staying home, in which case the division had been 40/42, the outcome would have been overdetermined. In that case, the outcome does not depend on any individual contributor (though it does depend on a *group* of individuals, obviously). This is odd, because in the 40/41 case, all are pivotal voters, and all are clearly causally responsible. They are causally responsible because for each we can say that the outcome depended on them. In the 40/42 case, none of them were causally responsible, because the outcome did not depend on their individual contribution. This should indeed strike us as odd, especially if an ascription of attributability responsibility and ultimately moral responsibility, depends on whether there were 41 or 42 people who voted in favour. So the voting case is almost always an overdetermination case. Only in the smallest scale voting cases, is the winning outcome unlikely to be overdetermined. So we can safely assume overdetermination holds. The relevant question is simply what kind of overdetermination occurs in the voting case.

3.3.3 Voting - a pre-emption case

I will now try to argue that the voting case, perhaps contrary to first appearances, can be described as a case of asymmetrical overdetermination, i.e. a pre-emption case. Recall, a pre-emption case is a case where there is a temporal discrepancy between the individual agents in question. This discrepancy lets us treat the counterfactual scenario as a different causal chain all together. Specifically, if Paul was part of the necessary set of votes⁴⁵, the case in which Paul fails to vote, is a different compound event. Because it is a different event altogether that obtains if he does not vote, the outcome in the scenario where Paul actually votes, thereby depends on Paul’s vote. Therefore, Paul is here causally responsible assuming the counterfactual analysis of causation. Indeed, it is not just Paul who

⁴⁵ i.e. the first x number of votes required to make the candidate win.

is causally responsible. Everyone who is part of the causally efficacious set is.⁴⁶ Specifically, everyone in the set of voters who were necessary for winning the day, is causally responsible. For each and every one of them, it is the case, that had they not voted, the total compound event of which the outcome is a part, would have been different. Therefore, they are each causally responsible for the event that actually obtains.⁴⁷

A large scale election proceeds very roughly along the following progression. People cast their vote in various places, at various times within some previously specified timeframe, usually on a single day, with variations depending on time zones. Further, some voters cast their vote days in advance and vote by mail. The votes are then counted locally, and the results are reported centrally, and the overall votes are tallied. At some point a clear winner has emerged, and the winner is then declared, usually before the last votes have been tallied. This is a rough picture, but the fine details are of less importance here. The important thing is that when we think of the voting scenario, we perceive of the voting result as an abstraction where all votes count equally, and are counted at the same instant. At first glance this seems to support the notion that this voting case is an example of symmetrical overdetermination. A somewhat anecdotal piece of evidence of this is that we tend not to rush to the voting booth in order to increase the likelihood of having a causal say and it is not the case that people rush to mail in their votes in an attempt to fill the causally efficacious set. Further, people voting in later time zones, do not tend to complain that they are being cheated, or that the process is unfair. However, contrary to first appearances,

⁴⁶ Note: the 'causally efficacious' set is the set of votes required to beat the threshold. If 100 people voted in favour and 50 vote against, then the causally efficacious set is 51, because that's what it took to win.

⁴⁷ E.g. imagine that voters A, B, C, D and E voted for the winner in temporal succession, who won by five to two votes. The outcome is then overdetermined, because it only required three votes to beat the threshold. The causally efficacious set was three, i.e. A, B, C. And the compound event was A, B and C caused the outcome. However, because there is a temporal discrepancy between them, it is a case of asymmetrical overdetermination. If e.g. B had not voted, then the causal chain would have been A, C and D. This is then a different causal chain, or a different compound event. A, B, and C causing an outcome is different event from A, C, and D causing that outcome. Therefore, the event that actually obtains, i.e. the one where A, B and C constitute the causally efficacious set, depends on each individual.

there is an actual matter of fact about which votes were actually cast first. All the votes have actual names attached to them, in the metaphorical sense of being cast by actual distinct individuals.⁴⁸ So we can e.g. imagine that Paul was the first of the voters who voted for the winning candidate, Susan was the second, Tim the third and so on and so forth. The main point here is that even though the results are presented as an abstract symmetrical threshold case, there is actually a temporal discrepancy between them as causally efficacious set filling causes. The outcome is not simply determined when the computer tallies the results. The outcome is determined when a sufficient number of votes are cast, votes that are individual contributions to the voting outcome. Therefore, the outcome is actually pre-empted, not overdetermined in the symmetrical sense.⁴⁹ This allows us to invoke Lewis' notion of fragility and separate counterfactual causal chains. Therefore, we can say, that the particular causal chain in which Paul voted, would not have obtained if Paul had not voted, as long as his vote was among those that filled the set. Therefore, Paul is causally responsible for the event which includes the voting outcome.

An obvious question is then whether all those who voted in favour of the winning candidate should be considered causally responsible. I would argue that the answer should be 'no'. The reason is that since we have to accept the temporal discrepancy between the voters, we also have to accept a temporal cut off point, where it is true that the winning threshold was secured. Beyond that point, i.e. beyond the pivotal voter (the voter who secured the outcome), no voter has any causal say in the outcome. They have no causal say because the event does not depend on them. If Ben (the pivotal voter +1) had not voted, the event would still have obtained, or put differently, the causal chain that actually obtained, would

⁴⁸ Of course voting is usually done anonymously, so we do not know which ballots represent which specific voter. But there is an actual fact of the matter, that every vote represents a unique individual.

⁴⁹ It is pre-empted, because if we imagine that the threshold was 41 votes, and Susan was the 41st voter, then it is true, that if she had not voted, then someone else, say, Allan, who in the actual sequence was the 42nd voter, would have been the 41st voters. So Susan pre-empted Allan's contribution.

still obtain. This leads us to conclude that any voter beyond the pivotal voter is a superfluous voter, and not causally responsible for the outcome. And if we assume that moral responsibility entails causal responsibility, at least in cases of positive agency⁵⁰, then we will also have to conclude that the voters who happen to cast their votes after the threshold has been secured, are not morally responsible for the voting outcome at all.

Though this account of causal responsibility seems initially promising, there are reasons for doubt. Firstly, the causal efficacy of the individual voter is implausibly strong. If we accept that the whole causal chain including the ultimate policy outcome (which is included in the compound event) is dependent on the vote of the individual, this seems like an incredibly significant event which can be attributed to the individual. It seems absurd to say that Paul brought about the voting outcome. In an ordinary voting case, say the recent UK “Brexit” referendum regarding whether to leave the EU, a single vote then plausibly has an effect of staggering amounts of utility or disutility (or let us at least assume this for the sake of argument). It seems that we are permitted to say that the individual caused the entire outcome, and is causally responsible for the entire outcome. We are allowed to say this because the whole compound event depends on our contribution. This will seem implausible to most, and it should leave us to doubt the particular account of causation.

Secondly, it will to many seem extremely unintuitive that the time a vote is cast is a plausible cut of point in terms of an ascription of attributability responsibility. In order to make this objection clear, imagine a small scale voting case.

Tribal voting:

A tribe deep in the rainforest is voting on whether to sacrifice the child of an innocent tourist, or set her free. Assume for completeness sake that they are all aware of what they are doing, and what the moral stakes are. Also, no one is coercing them or manipulating them. The voting is done

⁵⁰ As opposed to omission cases.

anonymously, where the voters put a piece of paper with their vote in a bowl, which is then counted at the end of the vote. Further, none of the voters had anything to do with capturing the tourist or the child, and none of them are going to do the actual deed. They simply vote, that is all. Someone else will count them, and carry out the deed. Assume then that three people vote against, and seven people vote in favour of sacrificing the child. In this case, the required votes to make the threshold were four. This then means that the three whose votes happened to be cast last, are not causally responsible for the outcome in light of asymmetrical overdetermination interpretation of the events.

According to Lewis' specification, we are allowed to perceive the outcome dependent on the first four voters, because the particular causal chain which actually obtains, would not have obtained if either of them had decided *not* to vote. The last three voters are however superfluous. Assuming we do not know which votes were actually the first four which were cast, most will still have the intuition that we can attribute the outcome of the innocent child getting sacrificed to all of them, even though we know that only a smaller subset could have been causally efficacious. This may of course in part be due to our ignorance of the *actual* sequence of the events. Perhaps we are just implicitly concluding that for each of the participants, the probability that *this* agent was part of the first four votes, is fairly likely, or at least more likely than unlikely, given the fact that the majority voted in favour. Perhaps it is not that we consider them attributability responsible per se, rather, we consider it *likely* that they are attributability responsible. However, I would imagine that even if we at some point after the voting was done, came to know with absolute certainty the actual sequence of events, this would not be sufficient to lead us to conclude that the outcome can only be attributed to the first four. In fact, it is plausible that our intuitions will be the same across both scenarios. I.e. our intuition will be that the outcome is attributable to all of them. In terms of a full ascription of moral responsibility, I would imagine that we also have the intuitions that they are all morally responsible. This seems to undermine that it is a commitment to counterfactual pre-emption that is at the heart of our intuitions of attributability.

If voting is an unclear case, we can instead imagine another simple firing squad example. Imagine then that the only difference between the shooters is a minute temporal, to the shooters unobservable, discrepancy. If we assume the firing squad consists of seven shooters, but it only takes four to kill the innocent child, I imagine almost everyone will share the intuition that the outcome is attributable to all of them. I imagine this is so regardless of the shooting order. But if causal responsibility is a necessary condition for attributability, and if the last three shooters are not causally responsible, that seems to contradict our strong intuition about the attributability responsibility of the shooters, and *a fortiori*, the voters, which is at least a serious blow for the idea that there is something special about the temporal order of the votes.

If our intuitions contradict the counterfactual analysis of causation in these cases, this may of course hint at the possibility that our intuitions need revision. However, as shown, these intuitions are quite resilient across cases. Alternatively, it may simply hint at the possibility that the counterfactual analysis, as here stated with regards to pre-emption, is simply wrong. It is too inclusive, because it allows us to attribute extremely significant outcomes to a single individual, even where she is one out of millions of contributors. Further, it is too exclusive, because it rules out attributing the outcome to individuals who should clearly be included. For at least these reasons, it seems that the pre-emption solution to the problem of overdetermination is unsatisfactory. The problem of symmetrical overdetermination still stands, and since we lack clear grounds for saying that the individual in all but the extremely unlikely case where the voting outcome hangs on a single vote is *causally* responsible, we also lack clear grounds for attributing the outcome to the individual. Of course, this does not rule out that the individual can be morally responsible, but it may rule out that she can be *directly* morally responsible.

3.4 Summing up the chapter

This chapter has proceeded in line with the strategy outlined in the previous chapter. Recall – the strategy was to assume that there were three individually

necessary and jointly sufficient conditions for direct attributability. In the beginning of this chapter I added credence to the assumption that causal responsibility is a natural way of tying an outcome to the agent, i.e. that causal responsibility is a necessary condition for direct attributability responsibility. Further, I assumed that a natural way of understanding 'cause', was in light of the counterfactual analysis of causation. Here an event causes another, if the latter depended on the former. I then argued that a problem with this assumption was that omissions are not clearly causes. Rather, they seem to be absences of causes. However, since it seems plausible that we can attribute outcomes to individuals in light of their omissions, this seemed to undermine causal responsibility as a necessary condition for attributability. The suggested solution to this problem was to treat omissions as a type of cause anyway, i.e. as quasi-causes. This however leads to an oversaturation of causes, which required another solution. The problem is that causal responsibility is simply too cheap, and an extreme range of omissions can be said to be the cause of something, even plainly ridiculous ones. If causal responsibility is supposed to tell us something important about when we can directly attribute an outcome to the individual, and hold her directly morally responsible, this is a problem. It is also a problem for the analysis. The reason is, if causal responsibility is not an individually necessary condition, and not part of a set of jointly sufficient conditions, then we cannot fully test the individual's direct attributability responsibility in the voting case by examining these conditions. Since the voting case is one where we lack clear intuitions going in, this is a problem. A response to this problem was to accept Thomson and Smiley's suggestion that we should import our normative commitments into our analysis of causal responsibility, in order to pick out the relevant omissions as "causes" in a particular case. The problem with this solution was however that it weakens the role of causal responsibility as a means for grounding attributability responsibility. Causal responsibility is now subjected to our moral intuitions, instead of being a way to uncover when someone is attributability responsible, and potentially morally responsible. The solution was then to set aside omissions, and merely examine whether positive actions can ground causal responsibility. The problem with this is then that we cannot fully test whether the voter is direct attributability responsible, i.e. we cannot in principle *rule out* her attributability responsibility by

ruling out her causal responsibility in cases of positive actions.⁵¹ Still, we do have enough resources to say that my voting action, or my voting omission, is an insufficient path for an ascription of causal responsibility. If I am causally responsible via an omission, it has to be in light of some other causal chain. Specifically, even if I do not vote, the absence of my vote would not have made a direct difference either.

I then examined whether the individual was causally responsible in the voting case specifically. If we understand “cause” in light of the counterfactual analysis of causation, there is then the problem of overdetermination. Specifically, the problem that the number of causes (voters) in play, entirely undermine the causal significance of any individual cause (voter), *taken individually*. The voting outcome does not depend on the actions of *any* individual except in the extremely unlikely scenario where the threshold is exactly met. Overdetermination is a significant challenge in cases like voting because of the particular threshold structure of such a case. In such a case, a certain number of causes are necessary to produce the outcome, where any number of causes before reaching the threshold is inefficacious, and any cause beyond this point is superfluous. A possible response to the challenge of overdetermination is to understand the whole causal chain that actually obtains, as dependent on the specific individual. This is possible if the relevant scenario constitutes a case of asymmetrical overdetermination, specifically a case of pre-emption. I argued that the voting case could be understood in this way, because there is actually a temporal discrepancy between the voters, and there is an actual set of votes which obtains which causes the outcome. The main problem with this solution is that it seems to entail that each

⁵¹ An example of such an omission could e.g. by me *not* fabricating an extremely convincing fake news story, disparaging the actually winning candidate, which would have resulted in a significant number of people changing their voting habits, to the extent that the actual winning candidate would have lost. My omission is then the quasi-cause of the actual winning candidate’s victory. Clearly the outcome of the candidate winning is not thereby attributable to be in light of this omission. However, the reason it is not, is not clearly that I lack causal responsibility for the outcome. More plausibly, it is because I am unaware of the significance of this omission, i.e. I did not know how to write and spread this particular news story. This pertains to fulfilment of the cognitive condition, which is the topic of the next chapter.

voter in the causally efficacious set, is then causally responsible for the *whole outcome*, which seems quite implausible in cases with millions of voters, and large amounts of utility or disutility at stake. We are e.g. forced to say of any individual, that she is causally responsible for the entire policy outcome. Another problem is that this understanding of causation implies that the succession of votes is important for whether we are causally responsible. If our vote is cast early in the day, we are causally responsible for the whole outcome. If it is cast very late in the day, we are not causally responsible for the outcome at all. This seems to place an implausible significance on the voting succession. Further, that the succession is relevant also contradict the behaviour of actual voters, who do not seem to care about the temporal sequence of the votes.^{52 53} Therefore, this is an unsatisfactory solution to the problem of overdetermination. Therefore, the problem of overdetermination still stands. If we accept that causal responsibility is a necessary condition for direct attributability, and if we accept that overdetermination undermines causal responsibility in the voting case, then we have to accept that the individual voter is not directly morally responsible for policy outcomes in light of her actions.⁵⁴

All in all, we lack clear grounds for holding the individual voter, as an individual, causally responsible for the voting outcome, at least in light of positive actions.

⁵² Unless e.g. their expressive reasons for voting are sufficiently strong to override their instrumental reasons for voting. However, we should assume that if the entire outcome was ascribable to the individual, she should indeed care a great amount about the succession of the votes, and people by and large simply do not.

⁵³ Of course, if it can be shown that this intuition is misguided, i.e. that we should care about the temporal sequence, then this would constitute a partial vindication of the causal responsibility via pre-emption. In that case.

⁵⁴ Unless we can give an account of some omission grounding our causal responsibility in such a case.

4 Cognitive condition

4.1 Introduction

In this chapter I will examine the cognitive condition, and explain in greater detail how it relates to direct attributability responsibility and moral responsibility. I will then apply the cognitive condition to the voting case and test whether it is plausible that the individual fulfils it. I will argue that she does not. As suggested in chapter 2, an outcome can only be attributed to the individual in the direct moral responsibility grounding way, if she fulfils a cognitive condition. It is plausible that this condition is fulfilled if the individual was aware⁵⁵ that the relevant outcome would follow from her action, or if she at least *ought* to have been aware. Further, it requires that she is aware of the significance of the outcome, or ought to have been aware. I will of course elaborate this below. Applied to the voting case, it concerns mainly whether she was aware that her vote contributed to the outcome, and whether she understood what a given candidate winning signified. Further, it also concerns whether she can be excused in cases of ignorance pertaining to both.

As noted, the preliminary strategy presented in chapter 2 was the following: Assume direct moral responsibility has two components, namely direct attributability and accountability. Then assume that upstream control, causal responsibility, and a cognitive condition, were individually necessary and jointly sufficient for direct attributability responsibility.⁵⁶ Then assume the citizen fulfils the upstream control condition, and test whether the individual citizen fulfilled the causal responsibility condition, and the cognitive condition. If she fails to fulfil one of them, then she is not directly attributability responsible for policy

⁵⁵ 'Awareness' is used in the factive sense here, such that if someone is aware of x, x is true. I will be using it interchangeably with "knowledge". There may be different epistemic standards with regards to each concept, but I will in this project assume there are not.

⁵⁶ Recall, if an individual is attributability responsibility for an outcome, then that outcome is representative of that individual as an agent, and it opens her up for a further ascription of accountability. If she is both attributability – and accountability responsible for an outcome, then she is morally responsible for that outcome.

outcomes, and thus not a candidate for direct moral responsibility. In the previous section, I cast doubt on whether the individual was causally responsible. Indeed, it seemed she was not.⁵⁷ In this section I will show that the individual does not fulfil the cognitive condition in the voting case either. This then rules out direct moral responsibility for policy outcomes.

With regards to the voting case specifically, the cognitive condition concerns broadly what the individual voter is aware of when performing the action that brings about the voting and policy outcome, and what she ought to have been aware of. As I will argue, this can be put into two related questions pertaining to the current part of the project of directly attributing policy outcomes to the individual voter:

- 1) Was the citizen aware of what the overall policy outcome of the election would be, and if not, is she then excused for her ignorance?
- 2) Was the citizen aware that she would make a relevant direct contribution to the policy outcome when she voted, and if not, is she excused for her ignorance if she actually did make a contribution?^{58 59}

Regarding question 1), this relates to the idea that we cannot directly attribute an outcome to a person, if she did not know that this outcome would be a consequence of her action or omission. When I donate money to a charitable foundation, I may have good reasons to expect that this money will be put to

⁵⁷ Though, it was not possible to entirely rule out that some unspecified omission constituted a quasi-cause, which could ground causal responsibility, at least in principle.

⁵⁸ This can also be stated as an omission. So again (assuming the same belief as before), If she fails to vote at all, or if she votes for a different candidate than the candidate who would – if she had won – lead to the morally optimal state of affairs, can she then be morally responsible for the voting outcome that obtains, if she did not believe her contribution would have made a difference to the ultimate outcome, if that outcome – at least in part – would have been different had she voted for the optimal candidate?

⁵⁹ Note that as far as I am aware, no one believes that the individual can be directly morally responsible for an outcome just because she believes she will bring it about, if it does not actually come about. She can be blameworthy for trying to bring it about, but there is not outcome in the actual world, which can be tied to her.

good use, e.g. because I have done extensive research. However, I may be entirely ignorant of the actual fact of the matter, namely that this money will end up in the pockets of some warlord, who will use it to cause large amounts of harm. If I was unaware of this fact, it seems inappropriate to attribute this outcome to me, unless I should somehow have known better. The same reasoning can be applied to the case of political participation. If I vote for a given political candidate because I believe this will lead to an overall good state of affairs, then I am not attributability responsible when it in fact leads to overall bad state of affairs, again, unless I should have known better when I voted.

Question 2) relates to the first, but instead of concerning whether I should have been aware of the significance of the policy outcome, it concerns whether I should have been aware that my action would *contribute* to the particular outcome. If I perform an action which I am unaware will lead to a given outcome, it is inappropriate to attribute the outcome to me, unless I somehow should have known better (which is often the case!). I will discuss this in detail, but it should seem plausible. If I lend my new phone to my friend, and it explodes in his hands due to an extremely unlikely battery malfunction, it seems inappropriate to attribute this outcome to me as expressive of who I am. At least if I was unaware that it would explode. We can e.g. imagine I had researched the phone prior to buying it, and were aware that there had not been reported any issues. This outcome is still something I am causally responsible for, because it depended on me lending him my phone. But, for it to be attributable to me in the attributability responsibility sense, I should plausibly also have been aware of this being a likely consequence of my actions. Relating to the current project, this question concerns whether the voting outcome and policy outcome depended on my vote. Assuming I actually make a contribution to the voting outcome, I may be excused for this, if I was unaware of this when I cast my vote, and if I am excused for my unawareness.

In order to respond to these questions, I will in 4.2 examine the necessary cognitive condition the agent needs to fulfil, in greater detail. I will continue to do this in section 4.3 where I will focus on answering when an agent is culpably ignorant, i.e. when she *should have been aware*. In section 4.3 I will also –

somewhat in brief - examine the voting case with regards to question 1). I will not give a full answer to this question, since doing that sufficiently is unfortunately outside the scope of this project, as it is mainly an empirical question whether the individual knows *enough*, and it will depend on the particular voting context. I will however highlight the significance of answering the question, and propose what is required for answering it. Following this, I will in 4.4 answer question 2). I will ultimately conclude that the individual usually fails to fulfil the cognitive condition in the voting scenario, because the individual is not aware that her contribution contributes to the outcome when casting her vote, even if it actually does contribute to it, further she is excused for this unawareness.

4.2 Cognitive condition

In this section and its subsections, I will elaborate and discuss why the cognitive condition is a necessary condition for direct attributability responsibility for an outcome.

Concerning full-fledged moral responsibility, many contemporary philosophers hold the view, that in order for someone to be morally responsible for an outcome, they need to have some awareness about the significance of that outcome, and that their actions will contribute to that outcome.⁶⁰ Fischer and Ravizza hold a similar position and contend that: "it would be odd to judge the driver [...]who backs his car out of his garage unaware that a tiny kitten is snoozing beneath the rear tire [...]morally responsible for the kitten's untimely death" (1998: 12). The idea is, even if we fulfil the causal responsibility requirement for attributability responsibility for an outcome, our lack of awareness regarding that which we are causally responsible for, undermines the attribution in the fuller sense of the outcome being tied to us as an agent, rather than as a thing. The reason it cannot be attributed to us in a way that potentially makes us morally responsible for it, is that our lack of awareness limits our capacity to engage with relevant aspects of the situation, even if our bodily movements produced the

⁶⁰ E.g. Levy, N. 2011, 2013, Zimmerman, M. 2008 occupy this position explicitly, but according to George Sher (2009) and Haji (2009), this view is widely accepted.

outcome (Levy 2011: 182). The facts are epistemically and thus cognitively unavailable to the agent, and the outcome can then not be an expression of our agency. This ties into moral responsibility, because in the case above, the morally relevant facts are not accessible to the driver. For the driver, it was simply not a scenario with clear moral relevance with regards to the life or death of the cat, so in a real sense, the scenario was not one with those moral implications.

That ignorance undermines our ability engage with certain features of a scenario, applies to both omissions and positive actions. Here is an illustrative example of an omission case with moral stakes involved.

Bike rider:

Imagine you are going for a bike ride in a local park. You have your earphones on, and you are listening to your favourite music. Unbeknownst to you, a small boy is swimming in a shallow pond only meters away from your bike trail, behind some trees. Unfortunately while swimming, he has an epileptic seizure, and loses his ability to stay afloat, and drowns in the shallow water. Saving him would have been an easy accomplishment for any adult person nearby, since the boy was small, and the water was only knee-deep. However, because you had your headphones on, you were unable to hear that he was in need.

If you had seen or heard him in need, and still just continued to drive on, it seems clear that we could have attributed the outcome of him dying to you, at least in part.⁶¹ However, because you were ignorant of the boy being in need, and because you had no good reasons to expect that anyone would be in need, we cannot attribute the outcome to you. This ignorance rules out you being accountable for his death, and thus your moral responsibility.

⁶¹ Even though most people distinguish intuitively between killing and letting die with regards to the severity of the act.

The above example was an omission⁶² case, but it is also clear in cases of positive actions that failure to fulfil the cognitive condition can undermine attributability responsibility, even in cases of clear causal responsibility. Below is an example of this.

Bridge collapse:

Imagine you drive over a bridge of seemingly high quality, but which has some hidden structural deficiency. Assume the bridge – unknown to any reasonable lay person - is close to meeting its incremental collapse threshold.⁶³ When you drive onto it, the combined weight of precisely all (say) 50 cars and trucks, plus the addition of your car, makes it collapse. Further, it would not have collapsed if you had not driven onto the bridge.

Now, in line with the previous chapter, you are causally responsible for the disaster (this is the case where you are exactly the pivotal *driver*)⁶⁴. However, we would not attribute the outcome to you aside from the causal impact you made. Surely you are not morally responsible for the collapse even though the bridge collapsing depended on your action of driving onto it. The reason you are not morally responsible, seems at least to be that you lack the relevant awareness about the strength of the bridge vs. your weight. It never enters your mind that there was any serious contingency which should have been avoided, and from the scenario, nothing indicates that you should have been aware of it. Though there is

⁶² As I noted in a footnote in chapter 2, some readers may have reservations about the term omissions in this context, and would simply refer to it as an absence. They may hold that for something to be an omission, the individual needs to fulfil some cognitive conditions, and actually be aware of what they are omitting to doing.

⁶³ E.g. it looked as fine as any bridge to your layman eyes, and everyone else was also using it at this point.

⁶⁴ Recall, if you are the pivotal driver, then you are the driver that fills the thresholds. In that case on the counterfactual analysis of causation, you are causally responsible for the collapse. If you had not driven onto the bridge, it would not have fallen (assuming no other driver would have filled the threshold instead). In cases of positive actions, as opposed to omissions, we can ascribe causal responsibility without a verdict of moral responsibility to sort through the possible causes, and highlight the morally relevant ones.

a sense in which you could have avoided the outcome (no one was forcing you), the fact that no red flags ever arose, means you never had a chance to take them into account when crossing the bridge. Therefore, we cannot attribute the outcome of the bridge collapsing to you. You could not have avoided the outcome, because you were unaware of certain relevant facts about the scenario which were necessary for making a morally relevant decision.

The main individuals in the two above examples are not attributability responsible for the outcomes. Seemingly, this is because they could not, in a real sense, have avoided the outcomes. Even so, there are reasons to doubt that it is merely because they could not have impacted the outcome due to the epistemic constraints of those scenarios. The fact that someone could not have avoided an outcome, or could not have done otherwise, does not clearly excuse someone or undermine their moral responsibility for a particular outcome, as Harry Frankfurt has famously showed (Frankfurt H. , 1969). His examples are well known so I will only present one fairly simple version of them:

Paul the wife killer:

Imagine Paul who wants to kill his wife. After a night of drinking with his unscrupulous friend Fred – a struggling neuroscientist – he professes his aim. Fred encourages Paul to kill her, with the hidden motive of testing a new action-monitoring/brain-control -device he has just invented. All *liquored-up*, Paul falls asleep at the bar, and Fred takes him outside to do a bit of back-alley brain-surgery, where he implants the device he just happens to have on him. Paul wakes up in his bed the next morning with a headache, but attributes it to a night of heavy drinking, and is none the wiser. The way the device works is that if Paul has already decided to do some previous specified action (or inaction/omission, had the scenario been different), the device starts to monitor whether the person changes their mind in regards to that action. If the person does change his or her mind, it will automatically block the decision, and thus make sure the person reverts to the previous course of action. The next day Paul kills his wife, and he never has the impulse to stray from that path. The machine

is never activated. However, even though Paul could not have done anything to avoid the outcome (since any impulse to stray would have been blocked in the counterfactual scenario), it seems clear that he is still morally responsible for having killed his wife, and any condemnation and punishment we would have directed at him in case he was not monitored, would indeed still be fitting in this case.

Frankfurt cases support the notion that our reasons for excusing and exculpating the driver and the bike-rider, may be something else than the mere fact that the individuals could not have avoided the outcomes, i.e. could not have done otherwise.⁶⁵ So we should look at the deeper reasons beyond the mere epistemic aspect (what the agent is or is not aware of) for why the fact that something blocks our access to relevant aspects of a scenario, seems to undermine attributability responsibility, and potentially moral responsibility. As I will show, the deeper reasons for why the agent is excused when she is unaware, is plausibly because her *mettle* is not tested. It is because the epistemic circumstances do not let the agent engage with the appropriate facts, e.g. moral facts.

4.2.1 Ignorance undermines reasons responsiveness

Ignorance does seem to excuse in certain cases. However, there is plausibly a deeper explanation for why ignorance mitigates or exculpates moral responsibility altogether, than the mere fact that it blocks us off from engaging with certain aspects of the environment. A plausible explanation has to do with the individual's reasons responsiveness (Fischer and Ravizza, e.g. 1998). The basic idea is, that if

⁶⁵ Note that various people have come to the aid of the *principle of alternative possibilities* as a condition for moral responsibility, most famously perhaps is the *flicker of freedom* response. The way a Frankfurt case is usually set up, there is some kind of tracking of the individual's deliberation. In this particular case, the device is set to track whether the agent has an impulse to change his mind. However, since such a tracking is necessary for the device to intervene, the example actually does rely on an ability to do otherwise. Changing his mind is a necessary condition for the activation of the device. So the Frankfurt example has not shown that we can be morally responsible even if we cannot do otherwise, because the ability to do *something*, otherwise, even if it is only having the impulse to diverge from the path, is built in to the example. (e.g. Fischer 2006: 10).

the individual does not have access to certain relevant features of the scenario, the individual cannot respond to those features. On this view, it is how we respond to reasons which grounds our moral responsibility for an outcome. The explanation for why we are not morally responsible in cases where we lack access to certain relevant considerations, is not just because we are ignorant, but in a deeper sense because our ignorance means we never get to consider whether to do the morally right or wrong thing. Indeed it still seems that attributability responsibility is an important part of an ascription of moral responsibility, but the deeper explanation why it is important, is because if we fail to be attributability responsible, we cannot make a choice whether to act morally or immorally. If e.g. the person on the bike had known that someone was in trouble, or if the driver knew the weight-limit of the bridge would be exceeded, they would then be open to an ascription of moral responsibility. They would be open to it, because in that case, they would be able to consider morally relevant reasons, and make a moral choice. If the person on the bike then decides *not* to help the child, or if the driver decides to drive over the bridge anyway, this tells us something about the kind of person they are. But in the cases above, their “mettle” so to speak, is never even tested. To reintroduce a substantive account of moral responsibility from chapter 2, their moral *ledger* has never been re-evaluated in light of how they responded to those reasons.

A way to look at this in more detail, is by examining J.M. Fischer's and M. Ravizza's (F&R) account of control, namely 'guidance control' along with their notion of reasons responsiveness. F&R concur with Frankfurt that the Frankfurt-examples show that moral responsibility does not require control in the sense of having the freedom to do otherwise, at least in some sense. The type of control which is undermined in Frankfurt's examples, which they believe is not necessary for moral responsibility, is what they call 'regulative control' (1998: 31). What regulative control is can be explained by an example, and by contrasting it with a type of control they argue *is* a necessary condition for moral responsibility, namely 'guidance control' (ibid). For an agent to have regulative control, it is required that the agent has access to genuine alternative possibilities. Specifically, the individual should have been able to have done otherwise, in the actual sequence of events

that obtained. However, assuming determinism, or more locally, that we are constrained in our course of actions by antecedent psychological conditions, it may be implausible that we could have done otherwise in most or perhaps any sufficiently similar situation. So if regulative control is a necessary condition for moral responsibility, this seems to undermine morally responsible in the relevant circumstances.

In contrast, F&R argue that the necessary notion of control for moral responsibility, is instead guidance control. In order for someone to have guidance control, their action has to be initiated by their own reasons responsive mechanism. What that particular mechanism is, is kept vague. It could e.g. be construed as certain mental faculties, but the important aspect is that whatever it is, it is the thing which is taken to initiate some action which the agent reasonably considers *her own* (1998: 39). She *takes* responsibility - so to speak – for who she considers herself to be (her mechanism), and the actions which are the exemplification of who she takes herself to be (1998: 241)⁶⁶. F&R distinguish between three types of reasons responsiveness: “weak”, “moderate” and “strong”, where they argue that moderate responsiveness is the appropriate requirement. They describe this as such: “A mechanism of kind K is moderately responsive to reason to the extent that, holding fixed the operation of a K-type mechanism, the agent would recognize reasons (some of which are moral) in such a way as to give rise to an understandable pattern (from the viewpoint of a third party who understands the agent's values and beliefs), and would react to at least one sufficient reason to do otherwise (in some possible scenario). That is, a mechanism is moderately responsive to reason insofar as it is “regularly” receptive to reasons (some of which are moral reasons), and at least

⁶⁶ The notion of “taking responsibility” is important for their account. It involves accepting certain dispositions, acts, beliefs, conceptions of self, tendencies and potentially more aspects of what is usually related to self-identification, as one's own. It is historical, and two people who are identical in the current time-slice, are not necessarily both morally responsible, because what they take responsibility for, is their historical selves. So if one person in the current time-slice is only identical to someone else, due to being the target of physical manipulation, then she is not the historical *owner* of who she is, and she has not taken responsibility for who she is, according to F&R. (1998: 243).

weakly reactive to reasons” (1998: 243-4). So instead of an ascription of moral responsibility relying on the agents ability to do otherwise in the sense of having regulative control, it is on F&R's account contingent on a sensitivity to reasons and a sensitivity to being able to act differently, had the scenario been relevantly different from the actual scenario. The “trick” here is that we can have guidance control, even if the universe is determined (and regulative control is undermined), while also satisfying our intuition that some type of control is a necessary condition for moral responsibility. F&R try to show us that the necessary type of control is guidance control. We had the freedom to act differently if the circumstances had been different. If we had not had our headphones on and actually heard that the boy was in trouble, and if we had known that we were exceeding the weight limit of the bridge, then we could have acted differently. However, if we *then* decided to ignore the child, and exceed the weight limit of the bridge, then we may be morally responsible for those outcomes.

Their account thus seems to explain why we would hold the individual morally responsible in the Frankfurt case where there is a counterfactual intervener, while incorporating a notion of control as a necessary condition for moral responsibility. The reason we should hold the individual morally responsible, is that in the actual sequence (the one where the intervener does not intervene), it is the agent's own mechanism (because he appropriately recognises it as his own) which initiates that action, and that mechanism is reasons responsive (let us assume). An upshot of this account is that it is in line with our intuitions, rules out people who e.g. have mental disorders, who have strong irrational compulsions, as candidates for moral responsibility, at least in circumstances which are somewhat akin to the actual circumstances. I.e. it rules out people who lack “upstream control”, the third condition for attributability introduced in chapter 2. The explanation for why they are ruled out, is that they would not have acted otherwise had the actual circumstances been *somewhat* different, precisely because they would not have been responsive to those relevant reasons. It also rules out some cases of those who are morally deficient, e.g. psychopath, because their account requires sensitivity to *moral* reasons. On the other hand, it still lets us incriminate ordinary people with an ordinarily functioning *mechanism*, i.e. ordinary reasonable people.

Fischer and Ravizza's compatibilist account highlights what seems an important part of the cognitive condition which most would find plausible, namely that it is necessary that the action is initiated by the agent herself, and that she is sufficiently responsive to reasons. Specifically, that she is not affected by outside manipulation. Further, and most importantly, it requires the ability to consider and weigh different (moral) alternatives with regards to initiating an action and making predictions about the outcomes of those actions, simply, it requires they are sufficiently responsive to relevant reasons. Though lack of awareness undermines attributability responsibility for an outcome, and thus seems to take the individual out of the running for being morally responsible for the outcome she produces, the underlying explanation plausibly is that it makes the relevant moral consideration inaccessible to the agent. Because these considerations were inaccessible to the driver in the bridge case and the bike-rider case, the agents are not morally responsible. But it is not just because they are ignorant; it is because – at least on F&R's account - their ignorance never triggers the relevant moral considerations for their reasons responsive mechanism to engage with. Again, their mettle is not tested.

We have now come full circle. The reason the individuals in the bike rider and the bridge case are not morally responsible, is because certain bits of information are unavailable to them. Had they been available, they would have been able to test their mettle, because in the particular circumstances where they have access to relevant information, they could have reflected upon them, and acted in light of them. If the bike rider then still fails to come to the aid of the boy, this tells us something about the bike-rider, namely that he would ignore the suffering of the child (which is clearly a morally wrong thing to do). Further, if the driver continued to drive over the bridge, knowing it would be fatal to many people, this shows us that he would be willing to bring about this suffering. Again, a necessary condition here is that the individual is responsive to reasons, because if she is not, the fact that she has access to certain facts about the situation is not enough to determine what kind of person she is. If the individual is not responsive to reasons, she is not sensitive to the relevant facts and cannot be judged by how she responds to

them. This rules out moral responsibility, and the fact that she failed to save the boy, or that she did drive over the bridge, does not allow us to attribute those moral outcomes of the boy drowning, and the bridge collapsing to her, i.e. it was not clearly her doing. Maybe she knew that the bridge would fall down, but she may simply have had a pathological compulsion to drive onto the bridge nonetheless. It would be outside her control, it would be involuntary, and she would thus not be morally responsible for the ensuing outcome.

4.2.2 Summing up so far

We can conclude that at least two things are necessary for fulfilment of the cognitive condition for attributing an outcome to the individual. 1), the individual's has to have access to the relevant facts. In the cases under examination, these are usually facts about whether her action will bring about some morally relevant outcome. 2), she has to have the capacity to engage with the information in the right way i.e. be responsive to reason (this is part of the upstream control condition from chapter 2)⁶⁷. This is e.g. not fulfilled if she was cognitively underdeveloped, if she was under the influence of some substance, or if she was under the influence of hypnotic suggestion. In the voting case specifically, we can assume 2) is fulfilled, i.e. the citizen is by and large reasons responsive. The question then is, whether she fulfils 1), i.e. whether she was aware that she by her voting action would bring about the voting outcome and whether she was aware of the significance, e.g. the moral significance, of the voting outcome, and the following policy obtaining.

Before considering the voting case in greater detail, there is however a further significant nuance to highlight. An important thing to note in the above is that a case can be made, that ignorance only excuses if the individual is non-culpable in her ignorance. In the next section I will examine when someone is culpably ignorant. Establishing this, will let me build on this and examine whether the voting citizen is sufficiently aware, and if not, plausibly culpable in her ignorance.

⁶⁷ From subsection 2.4.1

4.3 Awareness of policy outcomes and culpable ignorance

So far I have shown why ignorance can excuse the agent. It is because a consequence of ignorance is that it blocks the individual from responding to relevant morally significant reasons. However, ignorance does not always constitute an excuse. It may be the case that the agent is ignorant of certain morally relevant facts in the current instance, but if her ignorance can be traced back to a prior instance where she by her action or omission became responsible for her ignorance, she may plausibly still be morally responsible for the action or outcome in the current instance. In that case, her ignorance is plausibly *culpable*. In this section, I will examine when ignorance is *culpable* in this sense. Further, I shed light on the first question introduced at the introduction of this chapter. In the subsequent section (4.4) I will answer the second question from the introduction.

To recap, the question of whether an agent has a sufficient degree of awareness with regards to some outcome is an important matter, and for the case of moral responsibility for policy outcomes specifically, it relates – as noted at the beginning of the chapter - to two sets of questions pertaining to the voting scenario. Firstly, was the citizen aware of what the overall policy outcome of the election would be, and if not, is she then excused for her ignorance? Applied to a historically relevant case: was the citizen aware of Hitler's plans? Is she excused if she was unaware of the harms that would follow from his rise to power? Secondly, was the citizen aware that she would make a relevant direct contribution to the policy outcome when she voted, and if not, is she excused for her ignorance if she actually did make a contribution? Applied to the same historically relevant case: was the citizen aware that her vote would make a relevant difference to his rise to power, and if not, is she excused for her ignorance if she actually did make a relevant difference? If the citizen is excusably ignorant of the relevant harmful outcome, or if she is excusably ignorant of making a contribution to the outcome, then she is not directly morally responsible for policy outcomes. She is not directly morally responsible, because she fails to fulfil the cognitive condition for attributability responsibility, which is necessary for moral responsibility. In order to see whether she is excusably ignorant, we

have to look at when ignorance does not excuse, i.e. when someone is culpably ignorant. If e.g. the voter is ignorant of the significance of the outcome and of her direct contribution, but is culpably ignorant on one or both issues, then she may be morally responsible nonetheless, despite her ignorance.

4.3.1 Culpable ignorance

One way to understand culpable ignorance, is to understand it as attributability responsibility at one remove, i.e. whether someone is attributability responsible for an outcome they are ignorant of bringing about, because they are responsible for their own ignorance. On this account, a seeming case of ignorance being an excuse, is then undermined when the cause of the ignorance is attributable to the agent. In light of the current analysis, this would then be a matter of whether the agent was aware that she could have acquired awareness of what she is ignorant off. It could also concern the very special case of whether the agent had deliberately managed to cause her own ignorance (Aristotle, 2004: 46).⁶⁸

In order to elaborate the idea of attributability responsibility for one's ignorance, we can examine the following case.

Bush pilot:

Imagine that you are a pilot who flies people to remote wilderness areas. During one of your flights, you are caught in a violent storm, and your small plane goes down with several passengers on board. You are the lone survivor. Had you read the local weather report for the day (which you are aware of usually gives a good indication of whether there are good flying conditions), you would have known about the storm.

Even though you were caught by surprise by the storm due to your ignorance, your ignorance is seemingly attributable to you. If this is so, you are then

⁶⁸ E.g. someone tries to forget certain things, which would be relevant to some scenario in which they would have been morally responsible, if they had been aware of that of which they are now ignorant. By and large, it seems implausible that the general public can be accused of such a thing.

attributability responsible for the outcome you brought about in light of your ignorance. This then leaves you open for a further ascription of moral responsibility for the outcome of your passengers dying. Your responsibility was *seemingly* undermined with regards to the time-slice where you were surprised by the storm, but it was plausibly *not* undermined, because you ignored the news.

The reason your attributability responsibility was not undermined, we can assume, is that you were (let's simply agree) appropriately responsive to reasons, and you indeed had (we assume) good reasons to check the forecast. I.e. you were not blocked off from these reasons due to a deeper lying ignorance about the forecast having relevant information pertinent to your flight. Indeed, you knew the weather forecast was pertinent to your flight. Or in terms of proximate beliefs, you at least knew there was a *significant chance* that it contained pertinent information regarding the risks involved. A line of argument similar to this – namely that culpability can be “traced” back to an earlier culpable action, has also been defended by Holly Smith (1989), Gideon Rosen (2002), Neil Levy (2009), Michael Zimmermann (2008)⁶⁹ and the overall notion that ignorance can exculpate can be *traced* back to Aristotle (Nicomachean Ethics, Book III) as noted in chapter 2.

One thing that is missing from the explanation above pertains to the cost of acquiring said information. I specified in chapter 2 that with regards to a further ascription of accountability, the cost of defection should play a part. If I perform an action which foreseeably produces a significant harm, I may be morally responsible for it. But, if the cost of *not* performing it is very high, this will usually mitigate my accountability for performing it. Or, inversely, if I *fail* to perform an action (omission), in light of which some harm comes about, my accountability may be mitigated if the cost of performing the relevant action was high. The cost

⁶⁹ As Zimmermann acknowledges, this kind of culpable ignorance occurs much less often than we usually think. So we often misapply blame in cases on the basis of “he/she should have known”. (Zimmermann 2008: 178). The reason is, that in most cases where we think the person should have known better, they did not perceive any red flags. Thus, we will not be able to trace their ignorance back to an instance of *culpable ignorance*.

could e.g. be high if someone threatened to harm my family. So let us tie this consideration to the example from before. Let us assume that consulting the weather forecast was relatively costless, i.e. you did not have any other pressing – e.g. morally relevant - concerns which were incompatible with checking the forecast. Therefore, you are morally responsible for the crash, because you are culpably ignorant of the risk involved. In plain English – you should have known about the storm.

Compare the pilot to the voter who may be ignorant of the potential outcome of candidate A or B winning, but who has not spent more than a few minutes seeking pertinent information and deliberating over the relevant political issues. She is ignorant, but perhaps culpably ignorant. So assuming a moral-responsibility-grounding-causal-relationship obtains between the citizen and the voting outcome such that she is causally responsible. Here, even if the citizen is ignorant of the harm that obtains in part due to her action, she may still be morally responsible for it, or for contributing to it. The reason is that her failure to educate herself can make her culpably ignorant and thus responsible for the harm that obtains due to her culpable ignorance. I.e., she is not excused just because she did not know that her preferred candidate would start, say, and incredibly harmful unjust war, if it was reasonably something she should have known about. If it was relatively costless for her to obtain that information, she may be morally responsible for the outcome.

4.3.1.1 Was the citizen aware of what the overall policy outcome of the election would be, and if not, is she excused for her ignorance?

Whether citizens are ignorant of the effects of the policy outcomes of one candidate, party or policy winning, compared to the alternative, is an empirical question. As is the question of whether, in case of such ignorance, they are sufficiently aware of their ignorance to an extent where their failure to educate themselves plausibly makes them culpably ignorant.⁷⁰ Regarding the **former** (the awareness level of the citizen), there has been done a good deal of research on the ignorance of voters, and it does seem to be true that in many cases, the public

⁷⁰ It is of course also a theoretical question of what the standard they have to conform with is.

is woefully ignorant of the outcomes of one or the other options winning out (See e.g. Fishkin, J. 2004, and Brennan 2011: 163 for recent overviews). Though, of course, levels of ignorance/knowledge will vary across groups and individuals, and across issues.

The specific time around the second Iraq war is a case in point regarding the public's ignorance of what the policy outcomes entail.⁷¹ The Iraq war and aftermath has caused an enormous amount of death and the destruction, and though it was never up for a formal nationwide vote, it is plausible that public opinion did play a part causing the war, and it certainly played a role in its continuation.⁷² A large proportion of the US public were of the belief that the leader of Iraq, Saddam Hussein, had weapons of mass destruction, and based their support on this belief. With regards to policy outcomes, they presumably supported an outcome they perceived as the disarming of a dangerous dictator. However, since Saddam Hussein did not have these weapons, they were wrong about the policy outcome. Of course, there was no referendum on going to war, and the war was not advertised prior to George W. Bush winning the 2000 presidential election, so it is not a straight forward case attributing the war to the

⁷¹ Regarding the time around the start of the war, a Pew Research Center (2003) poll found that 76% of those queried favoured military action against Iraq if weapon-inspectors found Iraq to be hiding nuclear, biological or chemical weapons, while 63% were *directly* opposed military intervention if they could not find any. This is quite important, because 57% of the polled Americans were of the belief there had been found proof that Iraq were producing these weapons (Pew Research Center 2003). Further, public support in favour of the war on Iraq was at 72% at the time the invasion commences of March 2003 (Pew Research Centre 2008), so we can safely assume that a large proportion of those who favoured the war, based their support for the war on a false belief regarding the weapons capabilities of Iraq. This is thus important for what particular policy outcome they are supporting. If their support for the war depends on their beliefs about these capabilities, then they are presumably supporting what they perceive of as a disarmament policy. That is the policy outcome that they base their support on. If they supported the war *even* if Iraq did not have these capabilities, then they are presumably supporting another policy, e.g. a war to merely topple an evil dictator.

⁷² E.g. one study concluded that there had been more than 600,000 excess Iraqi deaths as a direct consequence of the Iraq war, where most of the deaths were due to violence (Burnham, Riyadh, Doocy, & Roberts, 2006).

American public. On the other hand, the Bush administration had to convince Congress to go to war, and those congressional representatives would presumably - this is all speculative - have an easier time justifying their support for the war, if the public were also in favour of it. Further, and more importantly, during the re-election of George W. Bush, his opponent campaigned (and lost) on a basis of withdrawing US troops from Iraq, so it is possible that this the outcome of this election depended on the false beliefs of a large proportion of the public. But this seems to suggest that their ignorance undermines, or would have undermined their moral responsibility. The question is then whether they are then culpably ignorant.

Whether parts of the US public were attributability responsible for their ignorance, is of course difficult to answer here. Answering that question depends on whether the relevant information was actually retrievable, what the cost of retrieving it was. It also ties into whether they were ignorant of their ignorance, which is an extremely difficult question to answer. I am unaware of any work done on this in relation to policy outcomes. Further, research seems to indicate that ignorance in general tends to increase one's ignorance of one's ignorance (Kruger & Dunning, 1999), so there are reasons to suspect that people would not be reliable respondents on questions pertaining to their ignorance, because they would be prone to overestimating their awareness of their ignorance.⁷³ A further question is of course whether people spend enough time acquiring quality information, and deliberate on that information sufficiently. Depending on this, they may be obligated to deliberate more than they already do, and thus they may be culpably ignorant for failing to do so. All of this of course still relies on whether they are responsible for bringing about the voting outcome and policy outcome itself, which is the most important question for this project.

Summing up this subsection. As noted, I will not be able to draw any clear conclusions on whether the individual is usually aware of the significance of policy outcomes. This is an empirical question which is beyond the scope of this project,

⁷³ People may be ignorant, but that does of course not necessarily mean that further deliberation would decrease their ignorance of these matters.

and one which is difficult to answer. It is particularly difficult because, even if we are able to conclude that the individual is ignorant of the actual policy outcome, this ignorance does not clearly exculpate the individual citizen. In order to answer the question of whether she is excused in light of her ignorance, it is also important to examine whether the citizen's ignorance is attributable to her, i.e. whether she is culpably ignorant.

4.3.1.2 A note on involuntarist moral responsibility and culpable ignorance

So far I have suggested that ignorance can be culpable, if the individual to some extent had control over her ignorance, i.e. if she was aware that she was ignorant, and was able to make a decision not to inform herself. However, not everyone agrees that ignorance needs be traced back to some conscious state about the awareness of one's ignorance about pertinent information, for someone to be responsible for something done out of that ignorance. Where the so far examined position can be considered a 'voluntarist' account of culpable ignorance, recently a number of writers have begun occupying an involuntarist⁷⁴ position in this regard. Involuntarists hold the view that individuals can be morally responsible in cases, even where it seems clear that they did not have any awareness of their ignorance, as long as the action, and the following outcome, is reflective of who they are as a person.

Proponents of the involuntarist view include at least Angela Smith (2008), Nomi Arpaly (2002), George Sher (2009), and Robert Adams (1985). The distinction between voluntarist and involuntarist could have an in principle effect on the application of a theory of responsibility to the current voting case. The reason is, it means that voters who are ignorant, who are not obviously in control of their ignorance in the sense of at some point being aware of it, could still – in principle – be morally responsible for outcomes which are a consequence of actions based in this ignorance. Robert Adams (1985) e.g. holds this view explicitly and writes that “the thesis that we are ethically accountable only for our voluntary actions and omissions must be rejected” (1985:3). On his view, we can e.g. be morally

⁷⁴ I am borrowing the voluntarist/involuntarist distinction from Holly Smith (2011).

responsible for our “involuntary sins” (1985: 3) in the broad sense that we can be blameworthy for our feelings of anger, jealousy, hatred and contempt, even when they are not within our voluntary control, either immediately, or when they are the result how we have deliberately acted to develop our character. Part of his reasons for holding this view, is based in the observation that it seems we may be obligated to make positive change to our character even when we did not cause it to be the way it is. This indicates that there is something wrong about the feelings and attitudes which were a product of that character in the first place. If they are not wrong, we would not have an obligation to make a positive adjustment to them it would seem. Further, the fact that we would praise someone for acting on the aim of changing their attitudes more if the motivation was conscientious, rather than if they did it – say – just to see if they could, indicates that what we are in part praising them for being conscientious. Further, more importantly, even if these character traits are clearly unchosen, the praise is still appropriate, perhaps suggesting that moral responsibility does not presuppose voluntary control. This is an in-principle distinction from the notion of voluntary control examined so far. If we accept this distinction, this does not necessitate, but opens the door to ascriptions of moral responsibility for involuntary acts, traits and outcomes. In that case, it opens the door for ascriptions of moral responsibility for outcomes which are in some way the product of ignorance.

Angela Smith writes in the vein of Adams and notes that: “what really matters in determining a person’s responsibility for some thing is whether that thing can be seen as indicative or expressive of her judgements, values, or normative commitments” (2008: 367). So with regards to the voting case, if Pauline votes for policy O, and O as a policy is close to a perfect match with regards to Pauline values, then Pauline can in principle be morally responsible for that policy (and the policy outcomes that follow) even if she was ignorant of the consequences of it. The reason is precisely that these policies are reflective of who she is. Angela Smith’s own example is of someone forgetting a friend’s birthday, without foreseeing it or intending to do so, but still being responsible for not calling her and giving her best wishes. The reasons I can be responsible here, is that what I am aware of (and fail to be aware of), and what I see as relevant of “my practical

deliberations" (2005: 270) can be reflective of who I am as a person, and a friend. The fact that I do not notice that it is my friend's birthday, may be because I do not value her as a friend, or because I am careless or perhaps have an egocentric conception of friendships. Forgetting her birthday is thus attributable to me, because it reflects who I am. This opens up for the same being applicable in the case of voting. I may be ignorant of the ensuing policy outcomes, and I may be ignorant of my ignorance. But if my ignorance, or my ignorance of my ignorance, is reflective of some fault which is expressed by this ignorance, I may still be responsible for the outcome due to this fault – even when it is an unchosen character trait. So if I vote for a harmful candidate, but I am reasonably ignorant of my ignorance that she will produce harmful policy, I may still be morally responsible for this if my ignorance of my ignorance is expressive of e.g. me lacking sufficient civic virtue in the sense of taking my electoral duty of informing myself seriously. The key thing to keep in mind here is that the agent in question did not have the prerequisite awareness of her ignorance, neither now, or at some previous time slice. However, she may still be morally responsible for an outcome brought about in light of this ignorance.

Nomi Arpaly argues for the same distinction between voluntarism and involuntarism, when she argues that people can be appraisable not just for their reflectively held views, but also for the type of person they really are (2002: 78). She examines the story of Huckleberry Finn who frees a slave, even though he himself under one construal considers this an immoral act. Finn is reflectively ignorant of the fact that this is really a good act (he thinks what he is doing is tantamount to stealing). However, we still praise him for his act though it is based on ignorance. It is here an ignorance of moral facts. His ignorance does not cut him off from praise, because he is seemingly still a good person. His action is expressive of him being a good person. Conversely, the person who catches a runaway slave would not necessarily be excused even though he thought he was doing good, if his catching the slave is indicative of his morally inferior character. The reason Arpaly gives for why we have this intuition, is that the moral fibre of the person is actually action guiding here, even though it is not consciously recognized by the person herself, i.e. even though Finn is ignorant of it. This can

be considered an in principle justification for attributing outcomes to someone, due to their unconscious mental states.

We can apply the above involuntarist considerations to the voting case. Part of the reason I vote for Hitler, may be because of deep seated hatred towards the Jewish population whom I would like to see harmed. It may also be because I failed to really listen to the actual content of his speeches, because of some character trait, e.g. my laziness. When people ask me, and even when I reflect on matters, I may respond and consciously believe that I support him due to him representing good conservative values and economic prosperity. I may even consciously disbelieve that he would ever do such a thing as harm the German and European Jewry⁷⁵. However, I do not challenge this view, or seek out relevant information. Or I am incapable of interpreting information in the relevant way, because deep down at an subconscious level, I do not want to find it, since that would perhaps force me to take a socially unacceptable stance on the issue, and I would have to force myself to accept that I held immoral views. Therefore, I am guilty of both being reflectively ignorant of my own motives, while I am also ignorant of being ignorant. I am then attributability responsible for both. Further, the policy outcome that follows from my voting action may also be attributable to me, assuming again I am also causally responsible. Importantly however, I am not in a clear sense in control of my belief, due to them being cognitively suppressed. Thus, if we accept that I lack control, while maintain that I can be attributable responsible for the above, then we seem to deny that moral responsibility requires awareness.

Summing up this subsection: It seems that ignorance (and ignorance of ignorance) does not always excuse attributability responsibility. There indeed seems to be in principle reasons to support this. Therefore, it is not necessarily enough to ascertain what people knew, and even whether they were ignorant of their ignorance. In the case of the German citizen who voted for Hitler, she may believe she is excused for her ignorance about the holocaust. However, whether that

⁷⁵ And people belonging to all the other groups that died in the holocaust.

individual is actually excused may depend on the person she is, and/or whether the person she is played a part in her failing to see the depth of her ignorance. The point is, any conclusion I will draw in light of whether the individual fulfils the cognitive condition, can be affected by whether the individual was ignorant, whether she was ignorant of her ignorance, and if so, whether this is excusable or not.

4.3.1.3 Summing up the sections on culpable ignorance

I have documented and argued that there are two interpretations of when a citizen fails to be excused for her ignorance of the significance of policy outcome she actually contributes to. On the voluntarist view, the individual is culpably ignorant if she had some awareness of her ignorance, and was aware that she could remedy that ignorance at a relatively low cost. On the involuntarist view, the agent may still be culpably ignorant, despite her ignorance, or her ignorance of her ignorance, if her ignorance, or her second order ignorance is e.g. due to some morally bad character trait.

In this section I have examined whether the individual can be excused for having false beliefs regarding the significance of the policy outcomes of the voting process. I will in the next section examine whether the individual is aware of her contribution to the policy outcome, and whether she is excused for her potential ignorance about this contribution. I will conclude that she is not plausibly aware of making a contribution when she does.

4.4 Cognitive conditions and voting

After having examined the condition for attributability responsibility, it is time to examine how the cognitive condition can be undermined with regards to something else. Specifically, how it can be undermined with regards to the individual's direct contribution to voting outcome. Recall, even if the individual fulfils the upstream control condition, and the causal responsibility condition for direct attributability responsibility, she also needs to fulfil the cognitive condition. I.e. she may be sufficiently free from constraints, and her bodily movements may actually bring about the outcome. But, if she is unaware of making a contribution

to the outcome, and is excused for her ignorance, we cannot attribute the outcome to her. If we cannot, this rules out direct moral responsibility for policy outcomes.

Assuming the citizen is fully aware of the significance of the voting outcome, i.e. the policy outcome which will be a consequence hereof, the question (introduce in the introduction of this chapter) is then:

Was the citizen aware that she would make a relevant direct contribution to the policy outcome when she voted, and if not, is she excused for her ignorance if she actually did make a contribution?

I will examine two different interpretations of how the individual can understand her contribution to the voting and policy outcome. I will examine whether she is aware of her contribution under one of those interpretations. Finally, I will examine whether she is culpably ignorant if she fails to be aware of her contribution. These interpretations mirror how her 'contribution' was understood in the chapter on causal responsibility (chapter 3). If one of these interpretations is indeed the correct account of how the voter actually perceives of her voting action, or ought to perceive of it. And, if we can assume that the citizen was not aware that she would make a significant contribution to the policy outcome by her voting. And, if she is excused for being unaware of this. - Then we can conclude that she fails to fulfil the cognitive condition for direct attributability responsibility for policy outcomes. This thereby undermines an ascription of direct moral responsibility. Instead, the outcome should in that case simply be considered a non-culpably unforeseen side effect of her voting action. As already discussed in the previous chapter, the causal relationship between the individual and the voting outcome is problematic. However, I will for the sake of argument assume that the agent is in fact causally responsible for the voting outcome, or part of it, in order to test the cognitive condition for attributability responsibility more clearly. The cognitive condition concerns awareness of certain facts, in this case facts about her causal contribution to the voting outcome. However, if she is not causally responsible for that outcome, then clearly she will not fulfil the

cognitive condition either on the direct account.⁷⁶ Since ‘awareness’ is factive, she cannot be aware of something happening, which does not actually happen. Therefore, I will assume she actually is causally responsible.

Regarding whether or how the voters ought to perceive their contribution: As I have already noted in chapter 2 and 3, voting is relevantly different from other types of collective ventures, in that it concerns a threshold, rather than a variable outcome which varies with every single contribution (Tuck 2008: 38). To recap this, in an example such as anthropogenic global warming, there is e.g. a variable outcome. Presumably every person who expels CO₂, does in fact contribute to worsening the greenhouse effect, however slightly. Thus in the global warming case, it is not simply a small chance that the agent contributes to the amount of greenhouse gas in the atmosphere. It is in fact a certainty (insofar as a person expels a net positive amount of CO₂ or alternative greenhouse gasses) that an agent will always be contributing something, though usually only a small amount.⁷⁷ From this an ascription of direct attributability responsibility for an outcome is fairly straightforward. The individual is at least attributability responsible for her specific contribution (assuming she is aware of what she is doing). In the voting case, which pertains to a determinate threshold⁷⁸, this is not obviously the case. Indeed, it seems that either one’s vote has an effect on the outcome, or it is entirely superfluous to that outcome.⁷⁹ As noted in the chapter 3

⁷⁶ This does not rule out that she fulfils the cognitive condition with regards to an indirect moral responsibility. Indeed, I will argue that she does in chapter 8.

⁷⁷ Everything equal of course. A person’s contribution may of course be neutral over a lifetime, or even negative. According to John Nolt, the average American Citizen is e.g. causally responsible for a significant amount of Greenhouse gas emission, to the extent where the consequences of this in terms of global warming, is equal to the death of two future people (Nolt, 2011). If he is right, that is then something we in principle can attribute to the individual citizen.

⁷⁸ A determinate threshold is the exact number of votes which is required to win relative to the total number of votes for the alternative candidate(s). In an election between two candidates A and B, insofar as it is true that A receives 2,000 votes, the threshold for B winning, is 2,001 votes. All votes above 2,001 votes are superfluous.

⁷⁹ Note, that the voter is of course responsible for the *act* of voting, as in putting a piece of paper in a box.

on causal responsibility, there were different ways of interpreting the individual's contribution to the outcome. However, the point remained, that one is either part of the threshold, or not. The outcome of a given candidate winning or losing does not vary a little by every vote, making the policy outcome better or worse. With regards to this, there are two possible ways the individual can make a *direct* contribution to the outcome. Again, I have already highlighted these in the chapter on causal responsibility, so this discussion will mirror it. Specifically, on the first interpretation, the individual can make a direct contribution by being the pivotal voter that exactly fills the threshold. On the second interpretation, the individual can make a direct contribution by simply being part of the causally efficacious set that fills the threshold. I will now examine whether it is plausible that the individual fulfils the cognitive conditions in either case.

Firstly, is it plausible that the individual fulfils the cognitive condition under the first interpretation? Is the individual aware that she is the pivotal voter when she casts her vote? Is she excused if she is unaware? As noted in chapter 3, this can be interpreted as the specific scenario where the outcome obtains, and where the threshold is met by one vote. In this scenario it is appropriate to claim, that the outcome depended on the individual voter (and all individual voters who voted for the winner). This happens in cases where the result without one's vote would be a real 50/50 split, that is, a situation where one's vote, given the votes already cast, would be sufficient to make one's side win. Now, assuming this unlikely scenario obtains, and the individual is actually causally responsible. Is the individual voter then plausibly aware of this at the moment when she casts her vote? In order to examine this, we can ask what the likelihood was that we (the individual) would be the pivotal voter. If it is too unlikely, it would be unreasonable for the individual to expect to make a contribution. If we take a country such as e.g. the United States of America, the likelihood of being the pivotal vote in a national election like voting for the presidency, is e.g. close to 1:100,000,000 (Riker and Ordeshook 1968).⁸⁰ This is extremely unlikely, clearly. Assuming the individual is aware of this likelihood, she almost never has reason to

⁸⁰ This has led some (e.g. Caplan 2006) to argue that when people vote, they are actually being irrational.

believe she will be the pivotal voter, at least in elections of this scale (it will be in the millions in most larger countries where it is a two candidate runoff election). In fact, it would seem entirely unreasonable to expect to make a contribution on this construal of the voting scenario. If we assume people are fairly reasonable, and have some fairly accurate grasp on the probability of them being pivotal, it is indeed difficult to imagine that they actually vote on the belief that it is likely that they will be the pivotal vote.⁸¹ If they lack the belief about something, they cannot be aware of it. Some empirical work has been done into this issue, and the conclusion is that people by and large actually do have a fairly accurate grasp on the probability that their vote will make a difference in this sense (Blais, A. 2000: 62-70). So it is extremely implausible that they vote with this aim in mind, and it would surely be unreasonable to do so. In the extreme unlikely scenario where they are causally responsible, it seems it should at best be seen as an unforeseen side-effect of their voting action. It is not attributable to the individual, because it is implausible that she was aware that the outcome depended on her in this way when she voted. She is surely also excused for being ignorant of the fact that she is the pivotal voter, because she knew it would be an extremely unlikely event and presumably, she could have done nothing to enlighten herself of the actual fact. Therefore, the voter is non-culpably ignorant in cases where they actually happen to be the pivotal voter. The outcome is not attributable to her.

An alternative and at first seemingly more plausible interpretation of how the individual contributes, is, as highlighted in chapter 3, in terms of asymmetrical overdetermination. Here the voters who actually filled the threshold, were those who were causally responsible for the outcome (section 3.3.1). On this interpretation, whether she fulfils the cognitive condition, concerns whether the

⁸¹ There are other theories for why people vote. E.g. there is the theory that we are motivated to a large extent by expressive reasons. (E.g. Brennan, G. and Lomasky, L. 1985: 195). On this account the larger part of why we vote has to do with expressing feelings such as (but not limited to): satisfaction for performing ones duty, the satisfaction from complying with the ethics of voting, satisfaction from affirming one's allegiance to the political system, party or candidate (Tuck 2008: 33) and even the satisfaction from the prestige obtained from "picking the winner" (Nadeau, R., & Cloutier, E. 1993).

individual is aware that she is part of the causally efficacious set, when she actually is. So, is the individual aware of this? Yes, presumably. This is the case, since we have reasons to expect that most ordinary voting scenarios are going to be fairly close, due to the vote seeking nature of political parties. So the voter is usually going to be part of this set, as long as her candidate wins. It is indeed a reasonable expectation that she would be part of the causally efficacious set. Therefore, it seems appropriate to say that she will be aware of this, even if it is not a certainty. It is a reasonably justified belief to hold when voting, assuming the preferred voting outcome actually obtains so it is not a stretch to claim she is aware of her contribution. Unfortunately, as argued, it is not a plausible account of contribution. She is not plausibly causally responsible in the asymmetrical overdetermination case, i.e. the pre-emption case. On this view, we are supposed to understand the whole policy outcome as attributable to the individual, because she caused it all, which is implausible. Further, it is also contradicted by actual behaviour, because we simply do not care whether we are the first, or the last to cast our vote. So while it is plausible that the voter is aware that she is part of the causally efficacious set, it is just not a plausible account of contribution that this in itself matters. If this was really something people cared about, they would rush to the voting booth to be one of the individuals who in part constituted the threshold, or at least, they would try not to be one of the last superfluous voters.

Summing up this short subsection. If we interpret her contribution as being the pivotal voter, then it is inappropriate to say that she is aware of her contribution. It would be an extremely unreasonable expectation, and she would clearly be excused for being ignorant of her contribution in this sense, when she casts her vote. The alternative interpretation of her contribution is however simply not a plausible account. She may be aware that she is part of the causally efficacious set, but we should not understand this as a direct contribution. Therefore, the individual does not fulfil the cognitive condition with regards to being non-culpably aware of her contribution to the outcome, even if she is causally responsible for the outcome by being the pivotal voter.

4.5 Summing up this chapter

The cognitive condition for direct attributability responsibility is fulfilled, if the individual is aware of the significance of the outcome she contributes to, and is aware that she actually contributes to it. A plausible explanation for why ignorance undermines her direct attributability responsibility and potentially direct moral responsibility, is because ignorance blocks off her ability to respond to relevant reasons. If it had not been blocked, she would have been able to test her mettle, i.e. make a choice whether to perform an immoral action or not. Even if she is unaware in the above noted sense, she may still fulfil the cognitive condition if she ought to have been aware of both when she performed her contributory action. Her ignorance may be *culpable*, if it can be traced back to an earlier point, where she had reason to enlighten herself, and could have done so at a relatively low cost. Alternatively, her ignorance is culpable if it was the product of a, in some cases unchosen, character flaw.

With regards to the voting case, the voter is morally responsible for the policy outcome, if she when voting, understands what that outcome is and the broader significance of it. As documented, this far from always the case. Further, she may not be excused for her ignorance, if she failed to enlighten herself to a sufficient degree, assuming this was relatively costless for her to do. Whether she usually is sufficiently enlightened, and does enough to inform herself, is difficult to answer. However, the important thing to note, is that an ascription of attributability responsibility depends on the individual's understanding in this sense. And in many cases where our intuition suggests the innocence of the voter, this may come down to the citizen's ignorance.

Lastly, even if she understands the significance of the policy outcome, she may still be excused for her contribution, if she was unaware of her contribution, assuming her ignorance is not culpable. In the case she is non-culpably unaware, it should be considered an unforeseen side effect of her voting action when she actually happens to make a direct causal contribution. As argued, out of the examined alternatives, there were no clear grounds for assuming that the voter should reasonably understand herself as making a significant direct contribution.

Mirroring her lack of causal responsibility, it is firstly unreasonable to expect to be the pivotal voter. Secondly, it may be reasonable to expect to be part of the causal efficacious set, i.e. the actual set of voters that brought about the outcome. However, it seems absurd that this should be the relevant understanding of contribution central to an ascription of moral responsibility, because it is simply an implausible notion of contribution. Presumably the individual voter does see herself as making *some* contribution to the voting outcome, when she votes and when her favoured candidate wins. But that notion of contribution is not captured by the direct account of moral responsibility. More plausibly, it is captured by an indirect account, as I will show in the latter part of this project.

All in all, because it does not seem possible to attribute the voting outcome, and thus the policy outcome to the individual directly, it seems that a direct account of moral responsibility is untenable. In the next chapter, I will sum up the first part of the project, and anticipate the last part, where I focus on indirect accounts of moral responsibility.

5 Rounding off all the previous chapters on direct Moral Responsibility

In this short chapter, I will sum up and round off the previous three chapters on direct moral responsibility. After this, I will briefly anticipate the next chapters on indirect accounts of moral responsibility.

5.1 Direct moral responsibility

In the last three chapters I examined whether the individual citizen is directly morally responsible for policy outcomes. I documented that such an ascription plausibly requires that the outcome can firstly be directly attributed to the agent, and secondly that the individual is accountable for that attributed outcome. If the individual is accountability responsible in light of her attributability responsibility for an outcome, she is then directly morally responsible for that outcome. The strategy I employed was then to set aside accountability responsibility for now,

and test whether the individual citizen was directly attributable responsible for the outcome. If she was not, that would then undermine her direct moral responsibility altogether, which would motivate examining indirect accounts of moral responsibility.

Regarding the conditions for attributability, these plausibly included three individually necessary and jointly sufficient conditions. Firstly, the individual has to be sufficiently unconstrained, i.e. she has to have upstream control. I did not engage with this condition in any great detail, aside from noting that the agent must at the very least not be physically forced or relevantly psychologically manipulated in order to fulfil it. Further, in order to fulfil this condition, the individual also needs to be free in a deeper sense pertaining to the discussion of free will and determinism. What kind of freedom the individual needs to have will depend on the particular theory one subscribes to, and has been left almost entirely untouched in this project. I will simply assume this condition fulfilled in the voting case. Secondly, the individual needs to actually have produced the outcome through her own action, i.e. the individual needs to be causally responsible. This *is* what makes direct attributability 'direct'. Thirdly, the individual needs to have been aware of her contribution to the outcome, i.e. fulfil a cognitive condition. I will sum up the chapters on these latter two conditions below.

In the chapter on the **causal responsibility** condition, I firstly discussed whether causal responsibility should appropriately be considered a necessary condition for direct attributability responsibility, i.e. for directly attributing an outcome to the individual. Beyond being obvious simply as a matter of the definition of "direct" attributability, I argued that causal responsibility is indeed a natural way of tying an outcome to the agent. I suggested that we should interpret causal responsibility in light of the counterfactual analysis of causation, where one event causes another, if the latter depended on the former. This however highlighted a problem for causal responsibility. The problem is that omissions seemed to undermine the claim that causal responsibility grounds, and is a necessary condition for, attributability. It seems clear that we can attribute certain

outcomes to individuals which obtain in light of omissions. However, it is not clear that omissions are causes under the counterfactual analysis. If we accept that we *can* attribute outcomes to individuals in light of omissions, this means that causal responsibility is not a necessary condition for attributability responsibility. That may not be a problem in itself, but it is a problem for the strategy of testing the individually necessary and jointly sufficient conditions for attributability.

A possible response is to accept that omissions are indeed not causes, but then hold that we can still include them within the general concept of causal responsibility as *quasi-causes*. Unfortunately this leads to the problem of causes overgeneralising which undermines causal responsibility as a grounding element of direct attributability responsibility and moral responsibility. There is always an extremely large number of omissions which quasi-caused any given outcome. A suggestion of how to narrow the field of *relevant* omissions, was to accept that relevant omissions which can constitute causal responsibility, in part are a matter of convention. Specifically, our clear intuitions of moral responsibility in an omission case, should then single out the relevant omissions which should count as causal responsibility. If I promise to cat-sit for my neighbour, and the cat dies of starvation, even though it may be true that all manner of things can be said to have *quasi-caused* the death of the cat, we have a clear intuition that *me* not feeding the cat, is the *morally relevant* omission, and thus the quasi cause that fulfils the causal responsibility condition. Therefore, it is also the relevant cause. The problem with this solution, is that causal responsibility is no longer a grounding claim, and thus no longer helpful in producing a verdict of moral responsibility in cases where our intuitions are unclear, as in e.g. the case of the moral responsibility for policy outcomes. I then proposed the strategy of leaving the problem of omissions to one side, and instead focus on cases of positive actions. In cases of positive actions, it seems to be clearer that causal responsibility is a necessary condition for moral responsibility and does seem to ground moral responsibility. If we assume that causal responsibility is a necessary condition for direct moral responsibility in cases of positive actions, this allowed for testing whether the individual was morally responsible for policy outcomes in light of her voting action. The important caveat to note however, is that we

cannot in principle rule out causal responsibility by merely focusing on positive actions. There could be some omission which would constitute causal responsibility, if we merely had strong intuitions of moral responsibility. We can only rule out causal responsibility for outcomes of positive actions.

With regards to ascribing causal responsibility to the voter, the most significant problem was overdetermination. Overdetermination is an inherent feature of the particular threshold structure of the voting case. Specifically, when more than a sufficient number of people vote for a given candidate, it seems true that the outcome did not depend on any individual voter. This is at least the case if we understand the voting case as a case of “symmetrical overdetermination”. When we do, all votes are considered to be cast simultaneously. The outcome does not depend at all on the contribution of the individual, taken as an individual, at least if more than enough vote for the relevant candidate. If voting is a case of symmetrical overdetermination, this seems to rule out her causal responsibility on the counterfactual analysis, except in the extremely unlikely cases where the individual voter is exactly the pivotal voter.

Another way to understand overdetermination, and a possible solution to the lack of casual responsibility, was in terms of asymmetrical overdetermination. Specifically, we may be allowed to understand the voting case as a *pre-emption* case. If we do, the temporal succession of the votes matter. In this case, those who cast their vote before the threshold was met, should then be considered causally responsible for the outcome, because the particular compound event in which they vote, depends on their individual contribution. Unfortunately, according to this interpretation of overdetermination, each individual who is part of the causally efficacious set, should be considered causally responsible for *the whole outcome*. In cases where there is a great deal at stake, e.g. in a typical national election, it seems entirely implausible that the individual is the cause of the whole outcome of the whole compound event. Further, understanding the voting case as a pre-emption case, goes against most people’s intuitions. E.g. it certainly contradicts how voters actually perceive of their voting action. In the pre-emption case, only those who actually fill the threshold are allowed to be

considered causally efficacious for the whole outcome. However, voters seemingly perceive of all their votes as counted equally, at the same instant (the same is true in structurally similar cases, firing squad cases). They do not rush to the voting booths to be among those who fill the causally efficacious set, even though in that case they would cause the entire event. We generally consider votes cast late in the day as important (or unimportant) as votes cast early in the day. So it does not seem like a plausible interpretation of the voting scenario.

Because the conclusion from interpreting the voting scenario in terms of asymmetrical overdetermination is implausible, it seems we have to accept that it is a case of symmetrical overdetermination instead. However, on this interpretation, the individual is not causally responsible aside from in the most unlikely of scenarios. I.e. only in case the threshold is exactly met, and our vote made the threshold. Therefore, the individual does not fulfil the causal responsibility condition in light of her positive voting action.

In chapter 4 I then focused on the **cognitive condition** for direct attributability responsibility. This was another individually necessary condition, which along with upstream control, and causal responsibility, is jointly sufficient for direct attributability responsibility. The cognitive condition concerned what the individual has to be aware of, and when she can be excused for lacking awareness. Specifically, in order for an individual to be attributability responsible for an outcome, she has to be aware that she contributes to that outcome, and she has to be aware of the significance of that outcome. Alternatively, she has to be culpably ignorant of both. The deeper reason why ignorance undermines attributability responsibility and potentially moral responsibility, is plausibly that it blocks the individual off from testing her mettle. When the individual is unaware that she is making a causal contribution to an outcome, we cannot say that this outcome is expressive of whom she is, because she never took a stand on whether to actually produce said outcome. This pertains both to producing the outcome, and the significance of the outcome.

Though ignorance may excuse, it plausibly does not *always* excuse and undermine attributability. One important exception is if our ignorance can be traced back to an earlier moment where we had control over our ignorance. Particularly, we may be ignorant about some outcome which will obtain due to our action. However, if we at an earlier point in time had a clear opportunity to enlighten ourselves about the potential hazards of that outcome obtaining, then our excuse may be undermined. If we indeed had the opportunity to enlighten ourselves (and were aware of that opportunity and our ignorance), and if the cost of doing so was sufficiently low, then we are plausibly culpably ignorant. In that case, the outcome of which we are ignorant can still be attributed to us. An alternative view is that we can be responsible for an outcome of our ignorance, even if we lacked control over that ignorance. If e.g. our ignorance could be tied to a particular problematic character trait, we may perhaps still attribute the outcome to us. We may e.g. have been unaware that we were ignorant about the significance of some outcome we contribute to, and we may indeed be ignorant of our ignorance. However, if the reason we are ignorant of our ignorance is because we are extremely arrogant, maybe our ignorance and our second order ignorance is not excusable. In that case, perhaps we can still attribute the outcome to us, in light of our arrogance.

With regards to the cognitive condition and the voting case: you may be ignorant of the significance of a given policy outcome. E.g. you may be ignorant of the fact that an overall “yes” vote on a referendum on whether to increase speed-limits on highways by 50%, will *actually* lead to an overall worse state of affairs compared to a “no” vote. You think a “yes” vote will lead to overall better state of affairs. So you are unaware of the facts of the matter. Further, imagine that the information indicating the actual consequence of the voting- and policy outcome was readily available, and accessible by a cursory internet search. Imagine also that you were either aware of pertinent information being readily available from trusted sources, or alternatively, it was available, but you were just too lazy to look into whether it was available. In this case, assuming you actually brought about the outcome through your voting action, we can plausibly attribute it to

you. Indeed, it is appropriate to do so even though you were ignorant of the *actual* significance of the outcome in question.

Aside from ignorance pertaining to the significance of the outcome itself, there is the further question of whether we are ignorant of actually making a contribution to the voting outcome in light of our voting action. Even if we are actually causally responsible for the voting outcome, we may still be excused for this contribution if we, again, were unaware of making it, while being excused for our ignorance. If we are ignorant, and excused, in the actual instance where we happen to contribute to the outcome, this “contribution” should instead be perceived as an unexpected side-effect of our voting action, and not an attributable contribution at all. I examined the two interpretations of ‘contribution’ from the chapter on causal responsibility, to see whether the individual could be said to be “aware” of making a contribution if she actually did. On one interpretation, it is simply unreasonable to expect to make a contribution, even if we actually happen to do so. This is the case where we are the pivotal voter. It is simply so unlikely that we actually are the pivotal voter that it is inappropriate to say that we have reason to believe we will be. If we do not believe it, we cannot be aware of it. Indeed, we should expect that the outcome would be symmetrically overdetermined, in which case we would fail to contribute. Alternatively, we should understand our contribution as pertaining to a case of asymmetrical overdetermination, i.e. a case where the temporal succession of the voters matter. Here, we were as noted causally responsible if our vote was part of the causally efficacious set. Indeed, it is plausible that we are aware that we will be part of the causally efficacious set. So we are aware of it, assuming that we reasonably expect our preferred candidate wins. Unfortunately, this is an implausible notion of contribution. This is in part due to the absurd conclusion that we are responsible for bringing about the whole policy outcome ourselves, because we were a necessary and sufficient condition for that outcome obtaining. It also contradicted how people tend to vote, where the succession of votes does not seem to matter. Therefore, it would be wrong to say that people are aware of *making a contribution* by being part of the causally efficacious set, because it is not plausibly a notion of contribution to at all.

Since causal responsibility and the cognitive condition are individually necessary for direct attributability responsibility, the individual voter is not responsible in this sense. Since we cannot attribute the outcome to the voter directly, she is not directly morally responsible either. For this reason, it will be necessary to see if we can give an *indirect* account of moral responsibility, which can incriminate the individual in the policy outcome of the state. I will do this in the latter part of this project.

5.2 Anticipating indirect accounts of moral responsibility

As noted in chapter 1 and 2, indirect accounts of moral responsibility can be distinguished from a direct account. They can be distinguished by not necessarily depending on a direct relationship between the agent and the outcome. Specifically, an indirect account does not require that the individual brings about, or necessarily makes a difference to, whether the outcome obtains, in order for that outcome, or part of it, to be attributed to her. What is then required is a matter of discussion which I will engage with in the latter part of this project, and it will depend on the specific account. However, as a very rough approximation, on an indirect account, the relevant outcome can be tied to the individual in light of her mere participation in a collective venture that brings about the relevant outcome. In chapters 6 and 7 I will examine different accounts of how citizens can incriminate themselves through their participation. I will ultimately conclude that none of them can be applied to the voting case satisfactorily. In chapter 8, I will establish an account which can be successfully applied to the voting case. In the last chapter (9), I will apply this account to the voting case, and conclude on the whole project.

The main reason for examining accounts of indirect moral responsibility is the plain fact the direct approach has turned out unconvincing in accounting for the citizen's moral responsibility in the title case. The individual is not morally responsible for her direct contribution to the voting outcome. However, examining the direct approach has been helpful in highlighting certain challenges an indirect account has to respond to, in order to ground a successful ascription of

moral responsibility. Firstly, quite obviously, it has to ultimately account for the moral responsibility of the individual in a way which does not rely on the outcome depending on the direct contribution of the individual. I.e., it must rely on something else than causal responsibility. Secondly, it will plausibly have to adhere to something like the cognitive condition. Even if her contribution is indirect, it has to be indicative of who she is. She has to be able to test her mettle. This can presumably only be accomplished if she is aware of her participation in some way, or at least ought to have been aware. Further, and related to these two challenges, a challenge will be accounting for gradability. I.e., if we cannot attribute the whole outcome to the individual directly, we seem to lack a clear yardstick for measuring the moral significance of the individual's contribution. An account of indirect moral responsibility needs to produce an alternative account of gradability. This last challenge will become more apparent in the next chapters.

Recall the structure specified in chapter 1: The next three chapters on indirect accounts will proceed as follows: In the next two chapters (6 and 7) I will examine different accounts of indirect moral responsibility. In the subsequent chapter (8), I will construct an account which can be successfully applied to the overall case, based on the preceding chapters. In chapter 9, the conclusion, I will apply this account to the voting case, where the specific policy outcome is a just war scenario. I will show that our intuitions suggest that the citizen is morally responsible in light of this account.

6 Indirect accounts of moral responsibility

6.1 Introduction

The overall aim of this project is to find the most plausible account of how the individual citizen can be morally responsible for policy outcomes. After having argued that a direct account of moral responsibility for policy outcomes is untenable, I will in this chapter, and the next, examine accounts which fit the category of 'indirect' accounts of moral responsibility. These accounts will prove to be more fruitful with regards to the overall aim. The chapters should be

understood as continuous, and have only been broken into two chapters in order to keep the discussion focused. The purpose of these two chapters is in part to examine whether any of these accounts can be successfully applied to the overall question, and in part to explore what is necessary for constructing a successful account of moral responsibility for policy outcomes. Though I show that they cannot successfully ground moral responsibility based on the individual's political participation, in doing so, I importantly also show why they fail. This allows me to draw some conclusions concerning what is required of an account of indirect moral responsibility, if it is going to justify holding the individual citizen morally responsible for policy outcomes to an extent which can justify a significant moral evaluation. The approach in this chapter and the next, will be to review fairly varied accounts. This will allow me to construct a successful account, based on the discussion of these, and the previous discussion of direct accounts. In chapter 8, I will then construct such an account, based on the conclusions drawn in the previous chapters. I will call this account a 'moral membership' account. In chapter 9 I will conclude on the entire project, and apply the moral membership account to the title case.

Before introducing the accounts I will examine, I need to make a note about the strategy so far employed. In the chapters on direct moral responsibility, I distinguished between direct attributability responsibility, and accountability responsibility. Here, the individual was morally responsible only if she fulfilled the conditions for both. I then suggested three individually necessary and jointly sufficient conditions for direct attributability responsibility. Specifically, these were: upstream control, causal responsibility and the cognitive condition (section 2.5). This motivated a natural strategy. Namely, test whether the individual fulfils these necessary conditions. If she does not, then we can rule out her direct attributability responsibility and thereby her moral responsibility. As noted, I simply assumed upstream control was fulfilled. I then examined whether the individual citizen fulfilled the other two conditions, which she did not. Unfortunately, this strategy is no longer applicable. Indirect accounts are more varied, and the particular conditions the individual has to fulfil will vary with each account. So, instead of specifying the necessary and sufficient conditions from the

outset, and testing whether the individual fulfils them, the strategy will instead be to examine and evaluate these accounts individually. Though the strategy from the previous chapters is no longer applicable, it will still be appropriate to distinguish between attributability responsibility and accountability responsibility. I.e., it will often be appropriate to distinguish between when an outcome can be attributed to the individual, and when the individual is then accountable for it, i.e. when the individual is an appropriate target of our reactive attitudes and other morally relevant sanctions (e.g. punishment). However, often I will simply write “responsibility”. When I do, I will be referring to *moral* responsibility specifically.

Regarding the different accounts I will be examining, I will focus on four fairly distinct accounts. The first type of account I will examine (section 6.2) is what I will refer to as a ‘simple reductionist account of collective responsibility’. This account by Sverdlik (1987), specifies that we can in be morally responsible for the outcome of a collective project, if that outcome was incorporated into our intention in the right way, when we performed our contributory action with regards to that project. This means, roughly, that if we are ten people who push someone’s car off a hillside, I am e.g. entirely responsible for the whole outcome if that was the outcome I intended. I am responsible for it, even if I was unable to bring about that outcome by myself, i.e. even if I was not causally responsible for it. On the other hand, if I only intended to push it as far as the edge, then that is the outcome I am responsible for, even if it happens to go over the edge. It is a reductionist account, because it reduces the question of attributability and moral responsibility, entirely to facts about the content of the individual’s intention. As I will argue, this is however also the weakness of the account. It is a weakness, because focusing entirely on the intention of the individual, makes this account unable to delimit the scope of the individual’s moral responsibility, to a plausible extent. Even in light of this, Sverdlik makes a convincing case that it is *mainly* facts about the individual’s mental states which ought to ground our moral responsibility in collective action cases.

The second type of account can be described as an ‘affective identification account’ (section 6.3). I will focus mainly on Abdel-Nour’s version of such an

account (2003). On his account, we can attribute an outcome to the individual, if an identificational link can be established between the individual and the thing that brings about the outcome. Such an identification is usually signified by the individual having feelings of pride in that outcome. This notion of identification establishes a potentially moral-responsibility-grounding continuity between the individual and the outcome. In the current case, if the individual citizen identifies with the state or with certain outcomes of the state, the individual can then potentially be morally responsible for these outcomes, even historical outcomes. I argue that this account is implausible, both because it is too inclusive and too exclusive. I can e.g. be attributability responsible for an outcome, even if I cannot help identifying with the particular outcome. Further, it is also too easy to distance myself from these outcomes since this requires *only* that I fail to identify with the outcomes. However, again, there is something plausible on this account. Namely, that certain facts about the individual aside from her causal responsibility, can create a continuity between the individual and a given outcome, even if it is not *mere* identification with an outcome, as Abdel-Nour suggests.

The third account (section 6.4) understands the citizen as being in a position to authorize the state, and the outcome the state brings about, where the citizen may then be responsible for those outcomes if she actually authorizes the state. I will focus mainly on Avia Pasternak's (2011, 2012) account because it ties most neatly into the current project, due to it being more strongly individualistic and participatory. On this account, the individual citizen can be said to authorize the state, if she is aware of her membership of this state, and further, if she fails to perform certain acts to distance herself from it. I will argue that her account, though otherwise convincing, fails to give an account of *membership* which is sufficient for moral responsibility through authorization. This is important, because it highlights that in order to be a member of something, the individual plausibly needs to do *something* to qualify for that membership in some way. I will build on this in chapter 8, where I - as noted - construct an account of moral membership, where insofar as the individual is a certain type of member of some collective project, she is then morally responsible for the outcome of that project.

In the whole of chapter 7, I will examine an account of joint action, specifically Kutz' complicity account (2000 and 2000b). On this account, insofar as something can appropriately be considered a joint action, the participating individuals can then be attributability responsible and accountability responsible for the outcome of that joint action, even when they fail to make a tangible contribution to the outcome itself. I argue that though his account is promising, it is too demanding to be applied to the case of individual moral responsibility for policy outcomes. However, I argue that his conditions for attributability and potential moral incrimination, can plausibly be loosened to fit the voting case, even if this means that it is no longer an account of joint action proper. I also argue that Kutz' account fails to specify a plausibly condition for when the individual *contributes* to the joint project, in a way which can potentially incriminate her in that project. This leads up to chapter 8, where most of that chapter concerns giving an account of the relevant contribution the individual has to make, in order to be incriminated in a group project, i.e. the contributory condition. The account I establish in chapter 8, is to a large extent a tweaked version of Kutz'.

6.2 Simple reductionist account of collective responsibility

On the simple reductionist account of collective responsibility, an individual is morally responsible for the outcome of a collective project, if the individual intended for that outcome when she performed her action in that project. The main proponent of this account is Steven Sverdlik (1987). His account is a response to the problems of ascribing moral responsibility in cases of collective action, where no single individual is clearly causally responsibility. I.e. it is a response to the problems for a direct account of moral responsibility, applied to cases of collective action. Though his account does initially seem promising, I will show that this account fails to adequately establish how the individual can be said to make a contribution in collective action cases.

The "problem of collective responsibility" which Sverdlik aims to solve, is: how can we hold multiple individuals responsible for a singular outcome, when more than enough contributors produce the outcome, where the outcome does not depend

on any single individual? So to take an example: if ten people push a car off a hillside, what outcome should we then attribute to each individual? One tenth of the outcome, or the whole outcome? Nothing? Are they individually responsible, or are they responsible as a group? If they are responsible as individuals, then what are they responsible for when the outcome in the causal sense did not depend on any single individual?⁸² In Sverdlik's view, the reason why these cases seem problematic, is due to a conflation of responsibility for actions and responsibility for outcomes. If we clarify the relationship between these two, the rest of the issues are easily solved, he claims.

The view of moral responsibility he subscribes to, is one where we are morally responsible for intentional actions, and in some cases for the *outcomes* of intentional actions. However, we can only be morally responsible for an outcome, if that outcome was already incorporated into our intentions when we performed the relevant action. 'Intention' for Sverdlik seems to roughly mean "foreseeability" of action (1987:66-7). Therefore, I intended to P, if I foresaw that P would obtain as a consequence of my action.⁸³ This is clearly reminiscent of the fulfilment of the cognitive condition from the discussion of direct moral responsibility, though it does not require any stringent notion of causal responsibility. Sverdlik's account of responsibility lets Sverdlik conclude that: "It is intention that is the solution to the problem of collective responsibility" (1987: 66). If we apply this solution to the case from earlier, what the ten people who push the car are then responsible for, is exactly dependent on the intention of the particular individual in question. I.e., it is what the individual foresees will be the consequence of his or her action. If one of them pushes the car, with the intention of leaning on it (with the foreseeable consequences that it would remain in place),

⁸² The example is of course quite artificial, because each contributor is at least responsible for lessening the load by some amount. So we should think of the example as something akin to a case of symmetrical overdetermination instead.

⁸³ He also seems to consider it inappropriate to say that wholly accidental outcomes (in the sense of unforeseeable) are outcomes the agent can be responsible for, since these as a matter of definition, are unintentional. This puts him in opposition to some of the involuntarist theorists discussed in chapter 4 on culpable ignorance.

unaware that the rest were trying to push it over the hillside (bear with me), then that is the outcome we can attribute to that individual on Sverdlik's account, and the outcome that individual is potentially accountable for. If instead one of them push the car with the intention of pushing the car off the hillside (and foresees this outcome), but only to give it a few dents, then that is the possible extent of what that individual is responsible for, and nothing more, even if the car completely explodes. And lastly, if one of them pushes the car with the intention of absolutely destroying the car, while foreseeing this outcome will be a result, then that is what he or she is responsible for. I.e. responsible for the full outcome of the cars destruction, regardless of the fact that the causal force necessary for that event to obtain, required more than his or her own strength. In the car pushing case just examined, they may all ten be morally responsible for this outcome (the car getting destroyed), but only if they all, individually, acted with the intention of producing the *same* outcome. I.e. only if the individually foresaw it.

In light of the above, it is clear why Sverdlik has a reductionist view of indirect moral responsibility. Collective responsibility is entirely reduced to the intention of the individual participant engaging in the group act. Responsibility is only collective in the sense of a simple aggregation of those who have the same outcome incorporated into their intention (1987:68). Therefore, Sverdlik concludes that there exists only one type of case where it is appropriate to refer to it as shared or collective responsibility. Specifically, in the case where two or more people act with the intention of producing *the same* result. It then follows that many seeming cases of collective responsibility, are on this account not actual cases of such. The reason is, that since genuine collective responsibility on his account only occurs when multiple agents act with the intention of producing the *same* outcome, there will obviously be many cases where the participants have diverging understandings of what their individual action will bring about. Further, even though it in some cases fits the description of "collective responsibility", due to the fact that several individuals act on the intention of bringing about the same outcome, the ascription of 'collective responsibility' in such cases, does not carry with it any special implications in terms of the moral

responsibility for each participant. The fact that I am part of an aggregation of individuals in this sense, does not make me anymore, or less, accountable for the outcome, compared to a situation where I was the only individual who foresaw that the particular outcome would be a product of our action. All individuals are still only individually responsible for their individual intentional action, and the outcome the individual foresaw would be a consequence of the action she performed.

A problem with his account is that it seems we can attribute outcomes to the individual, which she has not clearly caused or contributed to, but which obtain nonetheless. It simply requires that those outcomes are in fact incorporated into her intention when she performs her action. Because there is no required causal link between the individual and the outcome, we can in principle be morally responsible for outcomes which obtain, entirely unrelated to our actions. Sverdlik's claim is not that intention to produce outcome O is sufficient for attributability responsibility for O, because that would lead to the absurd consequence that we can be attributability responsible for outcomes which do not even obtain. However, he seems to mean that the intention to produce O, where O actually obtains, is sufficient for attributing O to the individual (presumably assuming the fulfilment of an upstream control condition). Recall again that intention means foreseeing. So if we foresee that O will obtain due to our action, and O actually obtains, then we are attributability responsible for O. This is however clearly problematic, since it does not incorporate any necessary condition that the outcome has to obtain due to *our* contribution. To show why this is problematic, imagine e.g. that I believe I have magical powers, which I believe allow me to bring about horrific events on the other side of the world, e.g. an earthquake. Assume then that I - the magician - perform (what I believe to be) the necessary incantation for this earthquake to happen. As a matter of (bad) luck, an earthquake actually happens, and thousands of people are harmed. It seems clear here, that even though this event was incorporated into my intention when I acted, I have done nothing to incriminate myself in that harm. It may be appropriate to consider me a terrible person, since I was willing to do what I thought would harm many people. My action was bad in *that* sense, and I am

presumably therefore blameworthy, to which Sverdlik will agree. But, it seems inappropriate to ultimately hold me morally responsible for the outcome in question. The reason it presumably is inappropriate, is because we lack some further story about how my action ought to tie to the outcome. But since Sverdlik does not, at least explicitly, seem to place any significance on any contributory relationship obtaining between the agent and the outcome, any lack of e.g. causal responsibility will not undermine an ascription of moral responsibility for the outcome in this case. The reason he does not want e.g. causation to play an explanatory role, is that this would cause trouble for the type of case he wants to resolve, namely the type of case where the outcome is causally overdetermined due to the number of agents involved, or cases in which the individual causal impact is insignificant. Thus, it seems we still need more for an ascription of responsibility in such cases.

Of course, he could stipulate that it should not only be a foreseeable outcome of our action, but a *reasonably* foreseeable outcome of our action. Some will then presumably claim that the *magician* from before, simply cannot reasonably foresee the outcome of the earthquake obtaining due to his magic. Therefore, we cannot attribute that outcome to the magician. The question is of course then, what should count as a reasonably foreseeable outcome? An obvious suggestion is that it is an outcome which we are justified in believing will obtain due to our action. Though this seems appropriate, it raises the question of what is meant by the “due to our action” part. It cannot be causal responsibility, because that concept will often not be applicable in cases of collective action, due to the problems highlighted in chapter 3. Further, Sverdlik explicitly rejects that I can be responsible for more than my own actions. So when Sverdlik writes that “more than one person can intend the same result. Therefore, more than one person can be responsible for the result even though each person is only responsible for his or her own action” (1987: 66) he fails to explain how my *individual* action, can be tied to a collective outcome, if it is not just the foreseeability of that outcome, but the reasonable foreseeability of how I *produce* that outcome. The problem is, clearly, that *I* do not produce it, *we* do, in some sense. Therefore, Sverdlik’s

account is lacking. He needs to give an account of the mechanism that ties the individual action to the outcome of the collective project.

Though it is lacking in this sense, there is something right about Sverdlik's reductionist account. Specifically that it is facts about the individual which grounds an ascription of moral responsibility for the outcome of a collective project. It fits well with this project, since this project exactly concerns individual responsibility. Further, it plausibly incorporates the cognitive condition from direct accounts. But, his account skips over actually accounting for the relevant notion of contribution. As will be a clear theme in the other indirect accounts under examination, it is accounting for the contributory relationship between the individual, and the outcome of a collective project, which poses the most significant challenge to an ascription of moral responsibility in collective action cases, and in the voting case in particular.

6.3 Affective identification and moral responsibility

A very different approach from those examined thus far, tries to show how we can attribute an outcome to the individual citizen, not through a direct contributory action, but through establishing a continuity between the individual and the nation that brings about the relevant outcome. This continuity is the proposed solution to the problem of the individual lacking causal efficacy over policy outcomes. Instead of showing how the individual can be causally responsible in light of her individual action directly, the strategy on this account is instead to explain how the individual can be a part of the nation itself, the nation which brings about the relevant outcomes. This then allows for attributing the outcome of the nation to the individual. Defenders of such an approach (e.g. Abdel-Nour, F. (2003), May, L. (1991), and to some extent Jaspers, C. 2000))⁸⁴ wants to explain how we can attribute acts of the nation to the individual by

⁸⁴ Jaspers' account is not one of moral responsibility for an outcome as such, but an account of *metaphysical guilt*, which attaches to those who are in some way associated with those who wrong someone, but does not make them accountable for those wrongs. Particularly in Jaspers work, the German people all share in this guilt because of the solidarity that existed among them in virtue of their belonging.

taking seriously the idea introduced in the introduction of this project, that the will of the people in some sense plays an important role in conferring moral responsibility on the agent. I will focus primarily on Abdel-Nour's account, since only he seems to be defending a sufficiently substantial notion of retrospective moral responsibility, suitable for the current project.⁸⁵ Abdel-Nour claims that individuals can be responsible for outcomes produced by the nation itself, both present and historical. The link between the individual and the nation, is the citizen's sense of national belonging. Specifically, it is the pride or guilt they feel in the accomplishments of the nation.

The reasoning Abdel-Nour employs is: "if by dint of her national belonging an individual can "win wars", "civilize barbarians," "build empires," and so on (which seems implied through the sense of pride in these among the individuals), is it not only logical to ask whether corresponding to these her imagined exploits she does not incur a responsibility for all of the bad states of affairs that these same actions have brought about?" (2003:702). The novel idea is that if it is entirely appropriate for me to feel pride in the accomplishments of the nation, and if this pride signifies that I in some normative sense am tied to these accomplishments, it seems – tentatively - to follow that I should also share in the responsibility of that nation. That which confers responsibility on the agent is not in a straightforward sense her deliberate actions. Instead, it is her identifying with the nation, i.e. other members which constitute the nation, who performs the actions she is then responsible for due to this national identificational link (2003: 696). Though he does not give a substantive account of the notion of responsibility he is subscribing to, he briefly notes that it relates to Bernard Williams minimal account of responsibility, which stipulates that to be responsible for something, the agent has to at least have brought about the relevant state of affairs (under some

⁸⁵ Actually, he writes as if it is a more substantial notion of retrospective responsibility he has in mind, but then seems to draw a conclusion pertaining to a significantly more limited notion of responsibility. I will however treat his account as an account of retrospective moral responsibility to see whether it can adequately be applied to the current case as such, nonetheless. Larry May's claim is still significantly more limited than Abdel-Nour's, May argues that citizens can only be morally *tainted* in these cases, but not fully fledged retrospectively morally responsible.

description), as well as fulfilling conditions similar to the cognitive conditions examined in chapter 4, i.e. some notion of awareness, and reasons responsiveness (or “state”, in Williams terminology – Williams, B. 1993: 55). The question is then how the individual actually *brings about* the relevant state of affairs by identifying with - and having a sense of pride in - the national accomplishments.

As noted, Abdel-Nour argues that the link is established when there exists a continuity between individual agent, and the nation which produces the outcome in question. This link is not meant to be understood as a physical link in the sense of causal responsibility introduced earlier. Rather, it is meant to be an alternative normative link. The link is supposed to be established through sentiments of pride in the particular accomplishments. So if I take pride in certain terrible events (e.g. the mass murder of a people) caused by the nation, and if it is appropriate for me to be proud of these events, this is presumably because I identify with the nation to some extent, and because of this, it is also appropriate to attribute these events to me. It is appropriate, because me identifying with the nation or its accomplishments makes the nation, even the historical nation, an extension of myself. Therefore, *I brought about* the events. This pride links me with those who performed those actions. In this sense, it seems my pride is evidence of, and my link to, moral responsibility.⁸⁶ How this continuity is exemplified is quite varied, and Abdel-Nour gives a range of examples: “From the perspective of those who participate in the national form of political identification, the nation is an association of which individual members are integral parts. Sometimes they go so far as to conceive of the nation via the metaphor of an organism. Even when they do not rely on such an image, they conceive of the nation as having a past, a

⁸⁶ Larry May (1991) has a comparable view, which is however less far reaching. When we are guilty of certain wrongs, we are morally tainted by these wrongs in a way distinct from being praise and blameworthy. Specifically, May focuses on our association with groups, which he believes can taint us in a – for moral responsibility – relevant sense, even if we are not contributing to the evil by our actions in an instrumental sense, and even if disassociating ourselves with the group, also lacks any effect on the harm. Our association is signified by our feelings of pride and guilt. One relinquishes oneself of this taint by disassociating oneself with a group.

future, a consciousness[...] They often [...] imagine it as having emotions (the nation can be humiliated, snubbed, or injured), a memory (it certainly knows how to hold grudges), and a will. As a result, those who take their national belonging seriously can meaningfully say things like “we have been injured,” “we won the war,” “we lost our country,” “we gained our independence,” “we made the desert bloom,” or “we shall prevail sooner or later.”” (2003: 698). Abdel-Nour considers it appropriate to talk of one’s membership in such ways, and he claims that these ways of speaking of our national bond, is evidence of there being an actual morally significant identificational link, where a person’s speaking of the accomplishments of the nation such as this, “solicits the admiration and approval of others, as if she had performed these actions herself” (2003: 702). This then, again, constitutes a continuity between the individual and the acts of the nation, which establishes a “type of causal link between them and the actions that brought about these achievements” (2003: 703).

Regarding the extent to which the individual may be held accountable in light of the attributed outcome, it is in Abdel-Nour’s view fairly limited. It does not justify punishment of the individuals in question. One reason is practical, where he is concerned about the consequences of retributive punishment for the individual’s in question, e.g. collectively punishing citizens. More principally, he rejects the notion of punishment because it would be “tantamount to criminalizing fantasies and feelings” (2003: 705), which has clear dystopian undertones. The idea is that “identification” is a sentiment, not an action with expected consequences. Instead, he suggests that the central notion of moral responsibility, implies that the individual ought to feel ‘guilt’ in her actions. This is of course, as Abdel-Nour hints at, problematic, since the pride that connected them to the outcome, will also constitute a strong psychological hurdle to feeling guilt in the morally relevant outcomes of the state one identifies with (2003: 709-10). For this reason, he acknowledges that ‘shame’ may be appropriate as a tool for bringing about the appropriate feelings of guilt, where we seemingly ought to adopt it in the national vocabulary directed at those who identify with these outcomes.

A problem with an affective account such as this, is that there may be too great pressure on the normative notion of the perceived link between the agent's identification and the outcome, to justify the necessary continuity required to say that I as an individual brought about the relevant outcome. Further, as John Parrish (2009: 125) notes about Abdel-Nour's account: while the mechanism seems to downplay the voluntariness with regards to being responsible (that we in some cases simply cannot help feeling pride), the opposite is also true. It seems too easy to distance ourselves from certain atrocities if we do not identify with them, but perhaps should (Neo-Nazi holocaust deniers come to mind). Beyond this, Abdel-Nour makes another claim which seems blatantly false: "national responsibility is actively incurred by individuals with every proud thought they have and every proud statement they make about the achievements of their nation. This, however, is also the limit of their national responsibility, which only extends to the actions that have historically brought about the objects of their national pride". The main contentious aspect of this is, that "statements" or expressive acts more generally, seem clearly insufficient to ground moral responsibility for the associated outcome. Surely, when an expressive act has e.g. foreseeable harmful effect, as when someone yells "fire!!!" in a crowded movie theatre, it can make the agent responsible for the ensuing chaos. However, in this case it is not the expressive act she is blameworthy for, but rather the instrumental implications of the expressive *act*. Simply expressing pride in something is clearly insufficient for being morally responsible for that something. This is even clearer when the expression is performed far removed, say in the isolated comfort of one's own home.

Further, identifying with something, and taking pride in something, seems at times an incidental psychological phenomenon, subject to rampant rationalization, which may or may not be, but often is, associated with performing an action resulting in some outcome. Further, we have a great capacity to identify with others, e.g. fictional characters, and even have a feeling of pride in their accomplishments. This can seem entirely appropriate, and not at all absurd, even when we acknowledge that they indeed are fictional. But this usually does not imply anything but a strictly sentimental link between us and the relevant

character, and we would never accept that *we* killed the dragon, or that *we* stole from the rich and gave to the poor. Part of the reason is of course that the outcome itself in this case is imaginary. But it raises the question of whether our national pride really implies a sufficiently strong notion of identification to say that *we* invaded Poland, other than in an entirely imaginary way which does not establish the relevant causal chain between us and the outcome. That being said, there is something to the idea that responsibility is associated with some notion of identification. E.g. on a Frankfurtian (Frankfurt, H. 1969) account of moral responsibility or other “real self views” (Wolf, S. 1990/1993), it is a requirement that the agent with regards to her own conception of herself, can assent to those actions (and their foreseeable outcomes) she is responsible for. The reason is, that only those actions are done freely (where some notion of a free action is a requirement for moral responsibility), in the compatibilist sense of the word. Or to use the vocabulary introduced in chapter 4, only those actions test *our* mettle. But as suggested earlier, this seems only a necessary condition for moral responsibility, and not a sufficient condition for moral responsibility as Abdel-Nour seems to suggest. So, an outcome should presumably be indicative of who we are, but more is surely required for moral responsibility. To underscore this, we can e.g. imagine a person who during the Nazi regime worked actively to undermine the holocaust, but who nonetheless could not help being psychologically disposed to feel some strong pride in those horrific outcomes. She would seemingly be morally responsible on the affective account, because the identificational link has here been established. However, this person could be the perfect Kantian moral agent (Kant, E. 1785/2011), who only does something out of duty itself in spite of her strong inclinations otherwise. Here it seems implausible that she should be accountable for simply having an inclination she cannot escape, even if she on some level wanted to. However, she *would* be morally responsible on Abdel-Nour’s account for those atrocities. And this seems extremely counterintuitive.

‘Identification’ may be tracking something of import to an ascription of moral responsibility. However, it is not clear that it in itself is significant to attributing an outcome to the individual, at least not to an extent that grounds a sufficiently

strong ascription of accountability. Presumably, it is important that the individual is not alienated from her actions in order to attribute the outcome of those actions to her. However, if identification should allow us to attribute outcomes which are not directly the consequence of the individual's actions to the individual, we plausibly need an account of some further mechanism for doing so. Abdel-Nour has not clearly provided such a mechanism, and it is thus not a suitable solution to the problem of how to incriminate individual citizens in policy outcomes. However, the idea that certain things can establish a continuity with a given project, or the outcome of that project is plausible. I will build on this further in chapter 8.

6.4 Authorization, group membership, and disassociation obligations

Another account of indirect moral responsibility, is the authorization account. Here it is not our notion of identification which makes us morally responsible. Rather, it is the idea that citizens as principals can authorise the state as an agent to do something in their name. The state is in this sense a representative of the citizen, and the citizen thereby becomes responsible for what the representative state does (policy outcomes). If what the state then does is within the scope of acting as a representative of the citizen, the citizen is then responsible for what the state does. Assuming we can indeed authorize someone else to act on our behalf in this way, the important question is then how we acquire the relevant position, which allows us to authorize the state, and which makes us responsible for what the state does. The suggestion in this section, is that we are in a position to authorize the state, if we have acquired a certain type of membership of the state. More generally, if we are a certain type of member of a group (e.g. a state), we also thereby stand for what the group stands for. If the group then specifically *is* the state, then we stand for the outcomes the state brings about.⁸⁷ . That is the

⁸⁷ This is e.g. the view of Parrish (Parrish, J. M. 2009), who grounds membership in light of the following idea: "Since only the state is capable of providing such goods as protection from harm, rule enforcement, and provision of key public goods, anyone who has signed on to the project of living the sorts of lives that require those goods may be seen from a normative point of view as authorizing the existence of the state that alone makes that project feasible" (2009: 131). This, and

notion of authorization at work here. When I am a member of a group in the right way, I thereby authorize the group to act. If I authorize the group to produce harms, then I am responsible for those harms.

Avia Pasternak (2012) gives an account of how individual citizens can share in the responsibility of the state in a way which makes the individual citizen liable for bearing the burdens for the wrongs of state, importantly, insofar as the citizen fails to voluntarily perform actions which clearly disassociate her from those wrongs (2012: 3).⁸⁸ Contrary to e.g. Abdel-Nour, part of the justification for this account is much more clearly participatory, which allows her to sidestep some of the contention from earlier, where one could be said to be responsible for what the state does, merely through identification.

the fact that citizens by and large do believe they have firm rights to those goods, indicate that they in fact sign on to the model, and further indicate - so the claim goes - that they perceive of themselves as the principals of a state as their agent, to an extent which: "commits us to prima facie obligations to take responsibility for the consequences of [the state's] actions" (2009:132). Parrish's account is however not clearly an account of retrospective moral responsibility, but rather an account of our prospective obligations. So it does not fit neatly with the current project. Marion Young sums up an account similar to Parrish', and states that "Our actions are conditioned by and contribute to institutions that affect distant others, and their actions contribute to the operation of institutions that affect us. Because our actions assume these others as a condition for our own actions, [...] we have made practical moral commitments to them by virtue of our actions" (2004: 371). Further, she writes, "these presuppositions of activity need not be present to an agent's consciousness in order to hold as assumptions. These relationships are objective." (ibid). So on this account, we need not take pride in the outcome, and alienation from the outcome is not an excuse even though: "It is not uncommon for agents to deny the connection to others their actions assume, but such efforts at bad faith are pragmatically inconsistent." (2004: 372). On her account, membership is then imposed on us when we go about our lives as citizen. It is imposed on us because our acts as citizens, default. Her account thereby lacks a clear voluntary aspect, and does not rely on the atomistic individual performing deliberate actions. So it does not fit well with the assumptions specified in the introduction of this project either.

⁸⁸ Examples could be where a given state e.g. suffers reparations for a wrong, such as those paid by the Germany to Israel after the Second World War (Pasternak's example 2012: 4), which are most likely passed on to the citizens via the tax burden.

Pasternak's authorisation account has two parts. Firstly, she stipulates that groups (e.g. states) in themselves, detached from the individuals who compose the groups, can be responsibility bearing moral agents in their own right. Secondly, she argues that individuals can be members of such groups, in a way where the responsibility of the group can be transmitted to them as members of the group. It can be transmitted to them as members, because their membership implies that they authorize the particular group outcome. The first part is supported by work done in corporate responsibility, where e.g. Peter French (1979) originally proposed that groups *as groups* can be morally responsibility bearing agents. They can be agents if they have the appropriate agential features, such as having identities which last over time, goals, internal decision making procedures, and the ability to produce outcomes based on these (in a sense, act). If a group fulfils conditions such as these, they may then be considered the right sort agents in their own right. They may be considered the sort of agent which can be an appropriate candidate for reactive attitudes, and sanctions (i.e. they can be accountable), which can be passed on to the individual group member. Since Pasternak does not herself try to defend the first part, but simply stipulates that groups can be morally responsible agents who can pass their liabilities onto the individuals who make up the group, my main focus will be on the second part. I.e., I will focus on when the individuals have the appropriate sort of group membership which can make them responsible for what the group does.

Pasternak's account of group membership is inspired by Christopher Kutz, whom I will return to in full in the next chapter. But, the basic principle of incrimination for what a group does (a group project), is that the individual can be incriminated in a group, when she has a 'participatory intention' to contribute to the relevant group. A 'participatory intention' broadly means the individual's: "intention to perform a certain act [which is] informed and rationalized by some goal, which is shared by individuals other than her" (2012: 8). When the individual shares such a participatory intention with others, they all then constitute a group. When they constitute a group, they can, as individual members, be responsible for the outcome of that group. Further, they can then be expected to incur the burdens for what the group does, i.e. for the group project, again, because they each

authorise the group to act. In the case where the group is the state, the individual members can then be responsible for the policy outcomes of the state. Noteworthy, it is not just the actual harm the individual group member causes herself, which allows us to attribute the outcome to that member. In principle, the individual can be more or less ineffective in bringing about any harm herself. She can be entirely causally inefficacious. However, because she may still appropriately be considered part of the group which does produce the harm, she may be accountable for what the group does, because she had a participatory intention which grounds authorization.

Pasternak stipulates that we can attribute the outcome of the state to the individual, in a way which makes the individual liable for what the group does, if she fulfils these four individually necessary conditions (assuming the group has corporate agency) for a **participatory intention**:

- 1) “The individual is a group member as a matter of fact, according to the group’s membership rules”;
 - 2) “The individual is reflectively aware of the fact she is a group member, and some of her actions are informed and rationalized by that fact”;
 - 3) “The individual is reflectively aware of - or at least can reasonably be expected to be reflectively aware of - the collective goals and activities of the group of which she is a member, and of ‘the instrumental relation of one’s part to the group act that is its end’. Alternatively (as would be the case in large complex groups like the state) the individual is reflectively aware of - or at least can reasonably be expected to be reflectively aware of - the fact that there are certain group activities of which she is ignorant.”;
 - 4) “the membership status is not imposed on the individual against her will”.
- (all quoted from 2012: 9-10)

In the case where the relevant group is the state, composed of its individual members, the individual citizen is then morally responsible for the policy outcomes of the state. Specifically, if the agent fulfils these conditions for having a participatory intention, then what the group agent brings about, can then be

appropriately transferred to the individual member, making her accountable for it.

Leaving aside whether citizens in general fulfil these conditions for a moment, are all these conditions plausibly necessary to incriminate the agent in the group (state) harm? The first thing one might note, is that it is not clear why a state has to be appropriately considered a corporate agent. As noted, I will not be discussing whether groups can have moral agency. However, it is simply not clear why it is required in this case. Pasternak seems to believe it is important for the group (or state) itself to be responsible, in order for there to be something (liability) which can be passed on to the citizens. However, I think this is unnecessarily demanding, and it limits the application of an account such as hers, since there will be many cases where the group clearly does not have group agency, but where our intuitions suggest moral incrimination of the individual nonetheless. As I will propose, I may intuitively share in the liability in some harm as long as I am be considered a normative part of a group of people which produce said harm, even when that group itself is clearly not a collective agent in Peter French' sense as described earlier. I.e. we can be incriminated in what a group does, even if it is more or less just a glorified disorganised 'mob' (no decision making structure, no identity over time, no internally defined goals) (French, 1979). I am going to elaborate this argument further in the next two chapters, but if I do my part in conjunction with the other group members, in bringing about some harm, and I foresee this harm may obtain as a consequence of the group, presumably that is intuitively sufficient for moral incrimination in that group project. As I will argue later, it seems sufficient that I contribute to the group act, in order to authorize what the group does, even if the group lacks the noted agential features.

Let me support this intuition by testing these conditions on a different case without a group agent, where our intuitions still suggest moral incrimination. Imagine e.g. an idealised example of what happened at the Roskilde Festival in 2000, where nine people were trampled and crushed to death under the weight of the crowd attending a concert. Assuming some individuals in the crowd fulfil

the four conditions above, it seems they – at least morally (not necessarily legally) – would be accountable to an extent that can ground strong moral condemnation for this terrible outcome. This seems to be the case even if the specific individual's contribution was overdetermined, and individually insufficient to do any harm. We can still attribute the outcome, or part of it to her. We can seemingly do so, even if the group of concert goers were simply a disorganized group of people, incapable of fulfilling the conditions for group agency.

Regarding condition 1) “The individual is a group member as a matter of fact, according to the group’s membership rules”. This condition seems to be met if she is considered part of the group by the internal standard of the festival, or the concert. This can e.g. be evident from her concert wristband, or perhaps by the testimony of the other festival goers. Or perhaps she is to any outside observer clearly a member of the particular group of people who are causing the event. Or perhaps to all of them.

Regarding condition 2) “The individual is reflectively aware of the fact she is a group member, and some of her actions are informed and rationalized by that fact” This condition can presumably be met if she is reflectively aware that she is part of a group of festival goers, or specifically, that she is part of the group which is causing the noted harm. E.g. she can see the people getting trampled, she reflects on this, and having done so, she does not relinquish her “group membership”, but keeps pushing along with the rest of the group, or at least does nothing to disassociate herself from it, say, by trying to walk away (assuming this had been possible).

Regarding condition 3) “The individual is reflectively aware of—or at least can reasonably be expected to be reflectively aware of—the collective goals and activities of the group of which she is a member, and of ‘the instrumental relation of one’s part to the group act that is its end’. Alternatively (as would be the case in large complex groups like the state) the individual is reflectively aware of—or at least can reasonably be expected to be reflectively aware of—the fact that there are certain group activities of which she is ignorant” This is met if she is aware, if

not of the group goals, since it is not necessarily a goal of the group, then at least of the plausible outcomes of the actions of the crowd of people. She is aware that if she, and the others keep pushing, the crowd is in all likelihood going to be harming certain people.

Lastly, recall condition 4) “the membership status is not imposed on the individual against her will”. This is met if she is actually able to disassociate herself from the group, so she can actually stop any time she wants without incurring any cost, and she knows this. Though the conditions have been mildly revised, it seems that if the person (along with other likeminded individuals) is responsible under the original conditions, she would also be responsible for the outcome of the group project here. So, assuming we share the intuition that the individuals who fulfil these conditions can be said to authorize the horrible outcome, and thus are accountable for it, this raises the question of why we really need to impose group agency on the group, in order for the individual to be morally responsible for the outcomes of the group. Our intuitions seem to suggest accountability, even if the group is not a corporate agent, but are just a disorganized group of people. Regarding the state specifically, it may be a group agent. However, it is not clear that we should require that it is, for the individual members to be responsible for the outcome of the group/state.

Ignoring group agency, the question is then: are these four conditions individually necessary and plausibly jointly sufficient for the relevant sort of participatory intention, which would confer responsibility on the agent for the outcomes of the state? We may have the intuition that the festival-goer is morally responsible for the harm, or part of it. However, that does not mean that the deeper moral principle behind this intuition is accurately captured by the noted conditions. First of all, we could question whether condition 1) is really necessary⁸⁹. If some mix-up or accident occurs where my official government record is never registered, or it vanishes, would that really be enough to absolve me of my responsibility? It seems an odd requirement that membership cannot be much more fluid notion as

⁸⁹ “1) The individual is a group member as a matter of fact, according to the group’s membership rules”

noted in the ‘Roskilde’ example. Obviously it may be *practically* necessary that the citizen has been registered somewhere, in order for the state to pass on certain obligations to the citizen in the sense of being able to actually tax the citizen. However, as a normative criterion for attributing the relevant outcome to the individual, it's difficult to see why it has to be a formal notion of membership. Presumably, it is sufficient that I have the participatory intention directed at the group project, or that it would at least be *reasonable* for me to consider myself a member under some more fluid condition, even if I was not a formal member. If we compare it to a bank heist, surely I may evade criminal sanctions if I sneak out the back door, and none of my compatriots *rat me out*. But, the fact that the police fail to register my membership of the group, does not make it any less appropriate to claim that I am morally implicated in the heist, even if I am able to evade prosecution.⁹⁰ Indeed, presumably something like condition 2 is sufficient to cover this contention.⁹¹ Presumably I only need to be reasonably aware of my contribution to some harm, or of my participation in the group which causes harm. In that case, I would not be absolved by a third party if the group for some reason did not recognize my membership status. Of course, being aware of something, implies that there is something to be aware of, i.e. awareness is factive. So what membership then amounts to still needs to be specified if this is going to replace the condition Pasternak specified. But if Pasternak's suggestion is that our awareness pertains to a piece of paper, or a computer file with our name on in some central administrations building, as condition 1) seems to imply, then that is not clearly what our awareness ought to pertain to. That is not what makes us the relevant sort of member for moral incrimination. It is simply too trivial for an ascription of moral responsibility. Regarding 3)⁹² it seems fairly appropriate,

⁹⁰ Pasternak actually does respond to this contention in a footnote, and states that when there is not a clear institutional notion of formal membership, then something like common knowledge, or mutual openness among the members will suffice for fulfilment of condition 1). I will discuss this in chapters 7 and 8, because there are reasons to think this is too demanding also.

⁹¹ “2) The individual is reflectively aware of the fact she is a group member, and some of her actions are informed and rationalized by that fact;”.

⁹² “3) The individual is reflectively aware of—or at least can reasonably be expected to be reflectively aware of—the collective goals and activities of the group of which she is a member, and of ‘the instrumental relation of one’s part to the group act that is its end’. Alternatively (as would be

though perhaps too inclusive. I.e., it seems more appropriate that I am aware of the probable outcome of the group, than its goals. A group can have certain goals, but be consistently unable to fulfil these goals. If I am aware that it is plausible that the group e.g. will fail to archive its goals, then it is not obvious that I can be incriminated in the group project by simply being aware of what their goals are assuming I fulfil the other conditions. Rather, what should be relevant is what I reasonably foresee will be the outcome of the group project, whether it is a goal of the group or not. Condition 4), the avoidance condition⁹³ is not obviously contentious either, and mirrors the aspect of direct moral responsibility, where the cost of defection has to be relatively low.

So, assuming I fulfil these conditions, am I then the kind of member of the state, which through authorization justifies the state in passing on its responsibility to me as a member? It is not obvious. The problem is that Pasternak has not adequately specified what the relevant kind of membership is. Overall, Pasternak's account gets us closer to the mark in terms of accounting for the responsibility of individuals by focusing on how we can incriminate ourselves in a wrong if we are members of a project, and fail to disassociate ourselves from a harm that is an outcome of that project. Though in its original form, it is not a sufficient account for clearly justifying an ascription of moral responsibility for policy outcomes. The main problematic condition is condition 1), which was supposed to provide the substantive specification for when the individual actually is a member of the group. The other conditions i.e. that the individual is aware of her membership, that the individual is aware of the outcomes of the group, and that the individual can leave the group, seem entirely appropriate. However, they lack a specification with regards to grounding the individual's indirect moral responsibility for the outcome of the group project (e.g. the outcome of the state). I.e. what are the conditions for the membership she is aware of having? We need an account of how membership is actually constituted. Again, this was

the case in large complex groups like the state) the individual is reflectively aware of—or at least can reasonably be expected to be reflectively aware of—the fact that there are certain group activities of which she is ignorant “.

⁹³ “4) Membership status is not imposed on the individual against her will”.

supposed to be done by condition 1), i.e. that a person was a member when she was a member according to the group's membership rules. But as noted, this is not sufficient for specifying membership. It is not, because if we have clear intuitions of moral liability in a case where we are a member according to the membership rules of the institution itself, this intuition is presumably not going to change due to e.g. a clerical error, where we e.g. evade formal membership due to some unique circumstances. This suggests that something else is doing the membership-grounding work. E.g., assume I am member of the state according to its own rules. Further, assume I am aware of this. Then one day a law is passed, which accidentally deems that every 1.80m tall left-handed man born in October to parents with a certain hair-colour, is no longer a citizen. This then undermines my citizenship. Surely, if I really was morally responsible before for what the state does, say if I voted for Hitler in full knowledge of the potential consequences, it seems odd that I can be excused in light of such a clear triviality. Therefore, as noted, I seem Pasternak's account is insufficient for grounding indirect moral responsibility.

The problem with Pasternak's account as I see it, is really the same problem that was highlighted in Sverdlik's account. We presumably need a specification for how/when the individual can be said to contribute to the group, or the group project. Surely, being a member, and being aware of one's membership is sufficient for being incriminated in what the group does. But that is not sufficiently interesting if we do not know what makes her a member to begin with.

I will continue this discussion in chapter 7, where I examine an account related to Pasternak's. In chapter 8, I will as noted develop a full-fledged account of how an individual becomes a moral member of a collective project in the first place, which will in part build on aspects of Pasternak's account. Particularly the notion that membership can ground moral responsibility.

6.5 Summing up this chapter

I have so far examined three quite distinct accounts of how individuals can be responsible for a collective project. In this chapter, I have examined accounts which try to specify how we can ascribe individual moral responsibility indirectly, where the individual is part of the project that brings about the outcome, but where the outcome does not necessarily depend on her direct contribution as an individual. Though they all seem to get *something* right, they are all problematic for various reasons. Sverdlik's simple reductionist account lacked an explanation for why something being incorporated into the agent's intention (was foreseeable by the agent) could incriminate the individual in a collective project. It presumably lacked a clear mechanism for how the individual can indirectly contribute to a group outcome. It specified that when the outcome of an action is incorporated into the intention of the individual, the individual can then be morally responsible for that outcome. However, Sverdlik did not specify the "action" part in any detail. So how an action incriminates the individual was unclear. Abdel-Nour's affective identification account, though perhaps being at the base of some of our ordinary intuitions about shared responsibility, seemed to lack a sufficiently strong mechanism for establishing the noted continuity between the individual and the acts of the state. Mere identification seemed insufficient. We can agree with him that feelings of pride in the outcomes of a collective project may hint that we see it as an imaginary extension of ourselves. We indeed may identify with the group or institution, and this can be entirely appropriate. But, in terms of something like an ascription of retrospective moral responsibility, we seem to need a more substantial continuity-grounding mechanism, aside for mere identification. Authorization accounts, such as Pasternak's, tries to account for such a mechanism. In Pasternak's case, it was roughly our formal membership, awareness of this membership, and the ability to opt out, which was sufficient for moral incrimination in the group project. But as noted, we still need some specification for what actually makes us a member.

In the next chapter I will proceed along the current trajectory, and examine a type of account which incorporates most aspects of the previous accounts, but brings

us closer to an account appropriate for the case of ascribing moral responsibility to the individual citizen for policy outcomes.

7 Joint action and complicity

7.1 Introduction

This chapter is a continuation of the previous, where I examine whether the individual citizen can be morally responsible for policy outcomes indirectly. In this chapter, I will be examining whether an account of ‘joint action’ can ground responsibility in the voting case. The main focus will be on Christopher Kutz’ thoroughly developed ‘complicity’ account of shared responsibility. I will present it, and then critique it, and then conclude that though promising, it fails because it is simply too demanding an account of shared responsibility to be applied to the voting case. In the next chapter I draw on his account, along with developments from the previous chapters, to construct an account of moral membership, which can give a positive response to the overall question.

Kutz’ account is a natural continuation of the previous chapters. Firstly, it is an explicit response to the failure of direct accounts of moral responsibility. Secondly, it tries to explain the continuity between the individual and the group, by accounting for how the individual becomes an inclusive author of what the group brings about. Contrary to the affective approach, it is not mere identification which grounds it, but rather active participation. Thirdly, just like with Sverdlik’s account, it is the intention to participate in the group project which in part incriminates the individual, though Kutz’ importantly couples intention with an account of how the individual also contributes to group (which then causes the outcome), something which is missing from Sverdlik’s account, and as argued, also missing from Pasternak’s account. Fourthly, this notion of contribution allows for his account not just to be limited to institutional contexts such as it was the case on Pasternak’s account, but also allows for application to less formal associations. However, as noted, his account of contribution needs to

be further developed in order to be successfully applied to the case of political participation.

Regarding joint action, as any paper or book on the topic will tell the reader, in everyday life, there are many clear examples of people performing actions, which can with varying degrees of aptness, be characterized as “joint”. These can e.g. range from walking together, dancing together, singing a duet, robbing a bank, and perhaps electing a president. Some of these joint actions bring about harmful outcomes, and tentatively, individuals acting jointly to produce harmful outcomes can in certain circumstances be morally responsible for the whole, or part of the whole outcome, just as they can be morally responsible for the consequences of individual actions. That is, they can *share* in the responsibility somehow. There has in recent years been a substantial amount of work done in this area in order to explain what it takes for discrete individuals to appropriately be considered as performing joint actions and having shared agency (central works include e.g. Bratman 1992, Tuomela & Miller, 1988, Searle 1990, Gilbert 1990, Kutz 2000, Kutz 2000b, Pettit & Schweikard 2006).⁹⁴ In the following I want to examine an account of shared responsibility that builds on the notion of a joint action. Specifically, Christopher Kutz’ strategy is as follows: he wants to give an account of how we can say that a number of people acted jointly, at least in what he calls a “limited” sense. He then wants to argue that this account, i.e. the conditions for joint

⁹⁴ For a distinction between individual actions, joint action and group agency, we may distinguish between them as Pettit and Schweikard (2006) do. A joint action is composed of individuals acting towards a shared goal who intends *that* “we” x together, and individual actions are those which fall short of this. Group agency is when the combined efforts of e.g. individual agents act in a way which creates an independent group agent which bear sufficient similarities with ordinary individual agents, to be considered separate agents in its own right. Such an agent appears e.g. when “members act jointly to set up certain common goals and to set up a procedure for identifying further goals on later occasions” (ibid: 33), while also setting up mechanisms for the joint agent to be responsive to reasons. As I will argue, shared responsibility requires neither joint action nor joint agency. The reason it does not require joint action, as I will argue, is it requires no strong sensitivity to the actions of others. It only requires that the individual is aware that her action contributes to some project, where the project is not necessarily an amalgamation of the participating individual’s conceptions of that project. I may in fact be the only individual who conceives of the project in a particular manner.

action, are also the conditions for shared responsibility. His claim is that when it is appropriate to say that a number of individuals perform a joint action, it is also appropriate to say that they are all individually morally responsible for the outcome of that joint action. Further, it is because they act jointly that they shared in the responsibility for the outcome. Importantly, because his account requires only the fulfilment of a limited notion of joint action, where e.g. participants are only required to have a vague conception of the other participants' intentions or plans, i.e. not full knowledge of the other participants action or contribution (Kutz, 2000: 67), it is very suitable for application towards the voting scenario. If voting is a joint action, it is, as I will argue, only in a very broad sense of the word.

Regarding this chapter, I will discuss his account on normative grounds in the sense that I will focus on seeing whether his account can sufficiently ground our intuitions of responsibility in collective contexts, one of these contexts being the large scale voting case. I will have less to say about whether Kutz' account is an adequate theory of joint action proper. The main reason for this focus is, as I will argue later, that though something being a joint action may be sufficient for moral responsibility, there are various cases, including the voting case, where the participants plausibly have shared moral responsibility, but where it seems clear that this responsibility is not grounded in it being a joint action proper. Whether or not Kutz' account accurately captures what a joint action is, may certainly be interesting in its own right, but that is not necessarily the most important aspect of when we can hold groups responsible or individuals responsible for what a group does. Further, focusing on the conditions for joint action may severely limit a given theory's application potential, particular to the current context. Regarding Kutz' complicity account, I will argue that it is insufficient for an ascription of moral responsibility for policy outcomes. However, I will argue that there are good reasons to loosen his account's necessary conditions, in order to include participants who fail to be incriminated on his account e.g. voters.

7.2 The complicity principle

Before introducing the Kutz' complicity account, I will briefly present Kutz' main test-case he discusses, in order to appropriately anchor the discussion and examine his account in detail. Kutz presents the case of the 1944 Dresden bombing and subsequent firestorm which – on his account⁹⁵ – needlessly killed more than 35,000 innocent civilians. In this particular allied air mission, approximately 8,000 crewmembers aboard 1,000 planes took part in the bombing raid (2000: 118), along with all the personnel who assisted, prepared, planned and commanded the mission on the ground. Whether these participants were aware of the moral significance of this attack is unclear, but let us assume that they were. The problem with accounting for the moral responsibility of each participant in this particular case – Kutz suggests - which he aims to solve, is that because of the large group of people who participated, the contributions of any single individual is clearly inconsequential to bringing about the outcome. He is repeating the concerns about overdetermination from chapter 4, and asserts that each individual airman was neither a necessary nor a sufficient condition for the harm to obtain, and thus did seemingly not cause the outcome. Further, each could seemingly rationally and honestly say to himself: “it doesn't matter one bit whether I drop my bombs on Dresden, or over the ocean. The outcome is already given; the same number of people will die”. Now, as noted in chapter 3, perhaps we are allowed to consider such cases a matter of asymmetrical overdetermination⁹⁶, both with regards to the time the individual airmen

⁹⁵ This is a controversial statement, e.g. Margaret Gilbert (2002) disputes the moral clarity of the case. However, for the sake of this project, we can hypothetically assume he was right without further concern.

⁹⁶ Recall, asymmetrical overdetermination is a case of overdetermination where certain discrepancies among the causes, allows us to understand the actual causal chain that obtains as unique, where if a cause had not been present, the outcome would have obtained, but it would have been a different event. Therefore, the event, or compound event, that would have obtained, would have been a different event. Therefore, the event that actually obtains, was dependent on the contribution of the individual contributor. In the voting case, the temporal discrepancy between the voters, is arguably, but not entirely intuitively, a sufficiently distinguishing factor in this way. In the Dresden-case, we may argue that from the perspective of the individual bombers, they did not drop their bombs at exactly the same time, in the same place, and that they did not causes the exact

contributed to the outcome, but also with regards to the particular pattern of destruction which is caused by their particular involvement, vs. a situation where they (individually) do not participate, so perhaps this case is not as problematic as first suggested. However, to the extent that we have an intuition that the individual participant is responsible for more than her own destruction, then Kutz is correct that this is indeed a problem case.⁹⁷ Kutz specifies that the reason why cases such as this are problematic is that they run contrary to two commonly held principles which Kutz claims are deeply rooted in common sense morality (2000: 116):

“Individual Difference Principle: (Basis) I am accountable for a harm only if what I have done made a difference to that harm’s occurrence. (Object) I am accountable only for the difference my action alone makes to the resulting state of affairs.” (2000: 116, emphasis his)

“Control Principle: (Basis) I am accountable for a harm’s occurrence only if I could control its occurrence, by producing or preventing it. (Object) I am accountable only for those harms over whose occurrence I had control.” (2000: 116-17, emphasis his)

As should be clear here, these principles seem to track what is at the heart of causal responsibility and the cognitive conditions for direct accounts of moral responsibility.⁹⁸ If these principles are true, this puts a clear limit on the scope of

same pattern of destruction that would have obtained. Therefore, if one of the bombers had failed to drop his payload, a different event would have obtained, making the event that actually obtains, dependent on the contribution of the individual.

⁹⁷ We can e.g. ask ourselves whether it would matter to the individual airman’s moral incrimination, whether his bomb had failed to detonate. And if it would matter, presumably we still have the intuition that he is more or less as responsible for the destruction as the other airmen. So we can presumably accept that there is at least some residual responsibility to go around which is not explained by his causal contribution alone.

⁹⁸ I.e. causal responsibility was an account of tangible difference-making. We were supposedly responsible for some outcome, because our action made a causal difference to it obtaining. The control principle is somewhat broader, because it also seems to include obligations to prevent.

the moral responsibility of the individual. Therefore, Kutz wants to argue for an alternative principle which is supposed to expand the scope of moral responsibility, namely the complicity principle. This principle is supposed to suggest how we can ground holding these individuals morally responsible in cases where a group causes some harm, but where the individuals within the group fails to make a direct contribution to the harm.

“The Complicity Principle: (Basis) I am accountable for what others do when I intentionally participate in the wrong they do or harm they cause. (Object) I am accountable for the harm or wrong we do together, independently of the actual difference I make” (2000:122, emphasis his)

The core idea here is that: “The notion of participation rather than causation is at the heart of both complicity and collective action” (2000: 134). This is in line with the intuition many will have that we can be morally responsible for more than what we ourselves do in a direct causal sense, i.e. we can be indirectly morally responsible for an outcome. Regarding the principle, note the ‘basis’ and ‘object’ elements. The basis of being accountable means the conditions someone has to fulfil in order to be morally responsible.⁹⁹ The object of being responsible is the thing (e.g. the states of affairs) which the agent is responsible for. Contrary to the Individual difference principle (IDP) and the control principle (CP) (or direct accounts of moral responsibility), the complicity principle, if it can successfully be defended, specifies that I can be responsible for what other agents do, or bring about, even when I fail to make a significant difference to what they do. The basis of the principle is – as with the IDP and CP – individualistic. It concerns the conditions the *individual* has to fulfil for an ascription of responsibility. But, the object of what we can be responsible for concerns a collective fact, namely a fact about the collective project we intentionally participate in. This is then supposed to indicate that this principle can bridge the acts of the individual, and the outcome of the group. Note, that the complicity principle does not exclude the

⁹⁹ Kutz refers to it as ‘accountability’, but it is clear that what he has in mind is retrospective moral responsibility, i.e. the same substantive notion of moral responsibility as is included in the topic of this project.

possibility that the individual may also be directly morally responsible due to something like the IDP and the CP, all while being complicit in the harm that others do. If I shoot someone during a riot, and the riot causes enormous amounts of harm, I may be both responsible according to something like the IDP and CP for the action and outcome of intentionally shooting and killing, while also being responsible for the outcome of the riot as a whole to some extent, due to something like the complicity principle. However, my incrimination in the group act does not depend on any ascription of direct moral responsibility (the IDP and CP), but instead on my intentional participation in the group project. If we take the air raid case, I may as an individual be responsible for my bombs killing someone whose death is entirely contingent on my action. However, my causal role, or lack thereof, with regards to the overall outcome of the air raid, is not discounted by my lack of efficacy with regards to the overall outcome. I may be both directly and indirectly morally responsible.

Regarding the scope of who can be morally responsible for a collective outcome, and to what extent, it is on Kutz' account extremely broad, as long as what I and everyone else does, can be characterized as a joint action, in Kutz' terms. Further, in e.g. the case of the Dresden bombing, what I am morally responsible for, is the ultimate outcome of 35,000 people dying, because that was the object of my action. This does however not necessarily imply, or so Kutz argues, that the appropriate moral evaluation of me, should be equal in harshness to a case where I had singlehandedly dropped a single bomb causing the same amount of harm (I will discuss this in section 7.5.). Different aspects of the situation may mitigate this.

7.3 Joint action

On Kutz' account, a joint action consists of individual agents who intentionally participate in a collective action while being suitably sensitive to the actions of the others (2000: 112). When we assign responsibility for what a group does, we look at the individuals who compose the group: "Because all collective action is reducible to individual action, accountability for collective harms can be nothing

more than the accountability of individuals who participate in collective acts” (ibid: 112).

For something to be a joint- or group action in the sense that makes them share in the moral responsibility for the relevant outcome, in line with the complicity principle, all the participating agents need to fulfil four individually necessary and jointly sufficient conditions (ibid: 75). The conditions are:

- Strategic responsiveness
- Goal sharing,
- Mutual openness,
- Having a participatory intention.

If they fulfil these conditions, then they can be said to act jointly, and if they can be said to act jointly, then the outcome that follows from that joint act, can appropriately be attributed to them all.

‘Strategic responsiveness’, refers to the idea that all participants’ intentions should be sensitive to the behaviour of the other participants, specifically their actions. This is required because “Jointly acting individuals do not merely act in parallel: each responds to what the others do and plan to do” (2000b: 5). This means that if Allan, Bernard and Charlie perform a joint action, then Allan needs to have some sensitivity towards the acts of Bernard and Charlie. E.g., if they are lifting a plank, they will need to have some idea about who will lift in which places, in order to know where to lift themselves, in order to accomplish the task. Though it concerns all individuals within the scope of the joint action, it is individualistic in the sense that it concerns the beliefs of the individuals in question (ibid: 68). Applied to the Dresden-case, a given pilot will fix his heading in relation to the supporting fighter planes and bombers around him. And/or perhaps he will wait till the maintenance staff has fuelled his plane before taking off. This condition does not demand that they require awareness about specifically who is doing what. Whether they need to actually respond to others actions, or just be ready to do so, is never specified clearly, but presumably an

action is not intuitively a joint action, if the participants are not *actually* engaged in some kind of coordination.

Regarding 'goal sharing', "Let us say two agents share a goal if there is at least one token activity or outcome, involving the actions of the other, whose performance or realization would satisfy the intentions of each" (2000b: 5). So Allan, Bernard and Charlie fulfil this condition, if they all have at least *some* goal in common, which they would all consider fulfilled, if one of the other participants completed it. It does not have to be the ultimate aim of the activity, it just requires that within the activity they are engaged in, *some* goal within that activity has to be shared between the participants, and considered fulfilled if one of them completes it. In that case, they are attaining *that* goal jointly. So if Allan, Bernard and Charlie intend to build a house, this condition is then satisfied if they all have a goal of building the house, where doing so satisfies each participant's goal when it is built. It does not matter whether it is Allan, Bernard or Charlie who actually fulfils that goal and builds, as long as one of them does. This condition is supposed to rule out accidental involvement to count as joint action. So if Dean "the crazy construction worker" had built the house in the middle of the night, before the group arrived at the construction site themselves, that person is only accidentally involved in the group act, and is not part of any joint action. Allan, Bernard and Charlie have not performed the joint action of building a house, and Dean had not performed it *jointly* either (only individually in this case). The reason is that the goal has been satisfied by someone else, who did not have an overlapping goal with the others. Neither we, nor he, included someone else who satisfied the conditions we all had set for completion of the task. In the bomber pilot's case, the goal sharing condition may refer to the overall goal of Dresden's destruction, and every inclusive member of the joint action will plausibly share *that* goal. If an accompanying airman slept through the mission briefing, and simply jumped into his airplane not knowing what the broad goal of the mission is, he does not share *that* goal. He may of course share other goals which would satisfy the condition still.

Regarding 'mutual openness', for something to be a joint action, all members must have an attitude of openness to the fact that the other participants potentially reacts to our choices. We must both be open to the others' activity, and them to ours in the sense that we welcome each other's participation in the project. Kutz seems to mean we must be open to the possibility that the other participants knows about our involvement, or could come to know of it. Further, if they came to know about it, they would be favourably disposed towards all of our involvement together. In accounts of joint action it is usually entirely central to such an account that there is something like common knowledge between the participants in order for it to appropriately be considered a joint action (see e.g. Tuomela, Gilbert, Bratman, Pettit & Schweikard)¹⁰⁰. But Kutz thinks actual common knowledge is too strong a requirement. The reasons is, his account also applies to large scale cases, where the participants only have a vague degree of awareness about who is actually involved. Therefore common knowledge would often be infeasible, even if common knowledge to many seems analytic to joint action proper. The less demanding condition of mutual openness however allows for the application of his account to situations where many people are participating in a joint project, but where some of the participants are unaware regarding some of the other participants' involvement, or to what extent they are involved. They nonetheless still need to have at least a positive attitude towards each other, if they came to know of the other's involvement. The bomber pilots in the previous example, are clearly not aware of the involvement of every other

¹⁰⁰ For something to be a group act proper (because this is on his view required for group intention which is a requirement for group action) Tuomela argues that it requires agents operate under what he refers to as "we-mode" (2010: 70), which incorporates shared reasons and beliefs, and a commitment to operate as part of the group, as opposed to I-mode, where an act can be entirely selfish, and done for private reasons. Beyond Tuomela, e.g. Michael Bratman stipulates the requirement for a mutual responsiveness of intention (1992), and Margaret Gilbert, stipulates that a joint commitment, which presupposes that the participating individuals are to some extent aware of, and open towards, such a commitment being made by all participants (2006, or 2013: CH1), is necessary for them to be appropriately referred to as what she calls a *plural subject*. But beyond these writers, it is generally assumed in the literature on joint action, that this notion of awareness is analytic to something being to be called a joint action proper (E.g. documented by Pettit, and Schweikard 2006).

individual, but they presumably still have some vague idea about the contribution of “the other airmen”, the “engineers on the ground”, the *brass* etc.

Concerning ‘participatory intentions’: Participatory intention is the self-regarding notion of acting with the intention of contributing to some groups act or outcome. By ‘self-regarding’, Kutz means that whether something is properly a joint action, is in part a matter of whether the individual herself intends for her own action to contribute to a collective outcome (2000: 74), while each participant who should be included in the joint act, share this intention, i.e. they have to overlap.¹⁰¹ He describes participatory intention thus: “an intention to do my part of a collective act, where my part is defined as the task [or role] I ought to perform if we are to be successful in realizing a shared goal” (ibid: 81). As is perhaps evident from that quote, participatory intention relates strongly to the other three conditions in its description. It incorporates the notion of goal, it presumably also includes strategic responsiveness, since my part will presumably be defined in relation to what others do (recall the example of a group lifting a plank). By referencing “we” it presumably also incorporates some openness towards the others. So it seems to include the other conditions in its description, at least implicitly. The participatory intention does not have to be conscious. The individual just have to have it. We can attribute intentions to agents: “on the basis of behavioural observations coupled with a general theory of human rationality”.¹⁰² This is important for Kutz, because he wants to be able to incriminate someone even if they are consciously alienated from the joint project they are a part of. They may not acknowledge that some goal was part of their conscious deliberation, but their actions can still be properly described as a contributory (ibid: 74), because their action can be

¹⁰¹ Kutz writes explicitly that they have to be fulfilled by each of the participants in an overlapping way. So I cannot be participating in a joint action if I do not have the same intention as the other participants, and vice versa. However: “So long as the members of a group overlap in the conception of the collective end to which they intentionally contribute, they act collectively, or jointly intentionally” (Kutz 2000b: 17).

¹⁰² He does not elaborate what ‘a theory of human rationality’ here refers to, but we can assume that rationality such a theory would also be functionalist non-intentionalist

explained by attributing a certain aim to the individual, based on our interpretation of their psychology.¹⁰³

Beyond this, there are two conditions which have to be satisfied here for something to be a participatory intention. Specifically, there needs to be an individual *role*, and a collective *goal* (ibid: 11). Specifically: “A participatory intention has two representational components, or sets of conditions of satisfaction: individual role and collective end” (ibid: 10). So I have a participatory intention, if I have some goal (which I am not necessarily conscious of), and the other participants share this goal. But I also need to fulfil an individual role in actually bringing about the outcome. By individual role he means: “the act an individual performs in order to foster a collective end” (ibid). By “foster” he seems to mean bringing about the outcome, or contributing to it. He then specifies that this role in bringing about the outcome can be either expressive, instrumental or normative. What he means by ‘expressive’, ‘instrumental’ and ‘normative’ is unfortunately not entirely clear, and he does not adequately elaborate them. We should think it means something different from standing in a causal relationship with the outcome or group act, since the particular lack of such a relationship was what his account of collective action was supposed to solve. Indeed, he acknowledges that this would lead to internal contradiction for his account (2000: 147). I will discuss the participatory intention condition in subsection 7.4.3 and here only briefly indicate some issues with Kutz’ understanding of it.

With regards to the role being ‘instrumental’ he writes: “Contributory relations might take instrumental form if what the agent does helps cause the collective outcome [...], or if the agent’s part is a constitutive element of the groups act” (2000: 82).¹⁰⁴ On the face of it, this may seem as if he is potentially reintroducing

¹⁰³ As already discussed in chapter 4, it is a divisive issue whether moral responsibility requires conscious awareness of the object of one’s responsibility, and presumably this also applies to an indirect account of moral responsibility.

¹⁰⁴ What he means by “constitutive part” could e.g. be someone moving in a certain way when dancing. What that means is unclear if it is not also a cause in a relevant sense. At any rate, the dancing example, which is Kutz’ own, gives the impression that there would not be a dance, if not

the IDP and the CP in the conception of role in the participatory intention. If this is the case, this is problematic, since participatory intention is a necessary condition for complicity, and he explicitly denies the IDP and CP as grounds for moral incrimination in his account. What could “instrumental” mean if not that? In elaborating what he means by ‘expressive’ role, he writes: “The relation might be expressive if by doing one’s part, one thereby exemplifies one’s membership in a group or participation in an activity, as when by voting I express my membership in a political community” (ibid: 82). It seems obvious that a merely expressive act is insufficient to incriminate anyone in any serious matter, especially if there is no instrumental relationship either (I will discuss this in detail in subsection 7.4.3). Further, for this project’s sake, the mention of the voting case is not helpful, since it is still unclear what about the voting action ties one to outcome of the group act.¹⁰⁵ With regards to what the normative role might be, he elaborates: “the relation might be normative if one performs one’s part because of norms internal to some group or institution that demand certain behaviour (I wear a dark suit as an IBM employee)” (ibid: 82). However, it is not clear why this is morally relevant, and it certainly seems that one can behave in this way without incriminating oneself in the harms of a group. He however elaborates: “Merely wearing appropriate clothing is not what constitutes my willing participation in [some organisation’s] corporate culture, but rather wearing [specific clothes] with the intention of being part of that culture” (ibid). But clearly that does not solve anything. It is simply restating that a certain *intention* is required for group membership. But, since what he is trying to do here is elaborate what a participatory intention even is, it is unhelpful to simply restate that it’s an intention to be part of some culture, as an example of a participatory intention. As

for the dance partner moving in this fashion, which seems to indicate that ‘constitutive’ means something like a necessary condition, which mirrors something like a cause in the simple counterfactual model of causation.

¹⁰⁵ It would help if he clarified why expressing one’s support for something is incriminating. E.g. yelling out the windows that I support slavery will not in any way tie me to that practice in a way that will make me responsible for it. At best, it will just make my neighbours look the other way when I meet them at the local grocery store.

noted, I will examine this central *participatory intention* condition in greater detail in subsection 7.4.3.

Summing up: I am morally responsible for the outcome of a joint action, if I fulfil four conditions. Firstly, I have to be strategically responsive, i.e. sensitive to the behaviour of the other participants. Secondly, I have to share a goal with the other participants, a goal which by the participants would be considered completed if either one of us brings it about. Thirdly, we have to be mutually open towards the potential participation of the other participants. We do not necessarily have to be aware of them, but if we came to be aware of them, we would be open towards their participation in a cooperative manner. Lastly, the participants have to have a participatory intention. This means that they should have a specific goal, and an intention to fulfil a role in bringing about that goal. This role can be either instrumental, expressive or normative.

7.4 Applying and critiquing Kutz' account

In this section I will examine whether the individual voter is complicit on Kutz' account. As I will show, his account is usually too demanding for this job (subsection 7.4.1). Therefore, I will in subsection 7.4.2 argue that there are plausible grounds for loosening some of the conditions for complicity. Further, in subsection 7.4.3, I will argue that one of the conditions for complicity, i.e. Kutz' "participatory intention" may be problematic for his account. It is a condition which is supposed to account for how we participate in, or contribute to, the group act, but it either fails at this job, or it seems to incorporate elements which Kutz himself has explicitly excluded from his account, i.e. the individual difference principle and the control principle. Doing this is problematic, because his account then fails to be a response to the problem that the individual fails to make a tangible contribution in certain collective action cases where we still want to hold them responsible.

7.4.1 Complicity and voting

The ordinary voter is not clearly complicit in light of her voting action. The problem is that there is no clear joint action, even in Kutz' limited sense, among the electorate as a whole, or even among the participants who vote for the same particular candidate. According to the complicity principle, I can incriminate myself in someone else's joint project, independent of making a difference to that project. I can incriminate myself in such a joint project, if I, along with others, fulfil the conditions for joint action.

First of all, we need to specify what the joint project is supposed to be in the voting case. Plausibly, it is not the whole voting venture. If it is a vote between good and evil, and I vote for good, then it seems absurd that I would be incriminated in evil winning, just because I vote in the election. That is clearly not a joint action proper, and certainly not one which would confer moral responsibility on me for evil winning. If my vote represents anything, it is a stand *against* evil. More plausibly, the joint project is *electing good*, or *electing evil*. So if I, along with others, happen to vote for evil, and evil wins, have we then incriminated ourselves in *that* particular project, and the following policy outcome according to Kutz' account? In order to see whether this is the case, I will have to examine whether the individual fulfils the conditions for complicity, i.e. the conditions for moral-responsibility-grounding joint action on Kutz' account. Recall, the individually necessary and jointly sufficient conditions were: "Strategic responsiveness", "goal sharing", "mutual openness", and "participatory intention".

So, does *strategic responsiveness* obtain between the voters who elect evil? This is unclear. In the Dresden bombing case, it is clear that the movement of plane A with regards to plane B, does require some responsiveness to the actions of the other. But in the voting case, I go to the voting booth, push my piece of paper through a slit in a box. That's it. There is no clear coordination between most voters, or readiness to respond to the actions of others. The inherent anonymity of the process ensures this. Perhaps if there is a queue of people at the voting place, I need to respond to the other voter's actions, but that is not clearly

participation in the joint action of electing evil. That is participation in the act of conducting an orderly voting ritual. So it would seem that ordinary voters in practice are too isolated and disunited to be strategically responsive to one another.

Regarding *goal sharing*, presumably many will have the same overall goal in mind, which would be considered satisfied if someone else fulfils it. In that case, they fulfil this condition. But we can at least imagine a voting case where this is not so. Imagine two citizens voting for the seemingly same outcome, but who have different, indeed wholly contradictory interpretations of what that outcome will be, if this candidate wins. I vote for Evil, because I believe his policies will decrease unemployment, and that is all *I* care about. Another voter votes for him, because she believes evil winning will increase unemployment, and that is all *she* cares about. In that case, there does not seem to be any shared goal between us. We seemingly have the same narrow goal in mind, “policy x winning the day”, but it is not the same policy we have in mind, so it is not the same goal we have in mind. They are actually, despite first appearances, working against each other. So it is not necessarily the case that this condition is fulfilled in the voting case. It depends on people’s conception of the project they are invested in. Some will vote for a candidate to support a single issue, and they do not know or care about what the politician stands for besides this. Others do not care about that issue at all, but still vote for the candidate, but then for entirely different reasons. So this condition is not always fulfilled, and it is certainly not fulfilled by all who vote for the same candidate.

Regarding mutual openness, this is not necessarily fulfilled either. Recall, this is Kutz’ substitution for the “common knowledge” condition for joint action. He plausibly claims that common knowledge is too stringent a requirement in many cases of shared responsibility. E.g. the bomber pilots do not necessarily know about who and how the other individuals participate in the venture. But they would presumably be open to the participation of the other individuals if they came to know about them, i.e. they would welcome them in a cooperative manner if they knew about them. In the voting case, mutual openness shares the

same problem as goal sharing. I may vote for evil, because I believe that candidate will benefit my social class, the upper class, to the detriment of the lower class. Another voter from the lower class, votes for evil, because he believes evil will benefit the lower class to the detriment of the upper class. If he is right, then I would not be open towards his participation. If I am right, he will not be open towards my participation. We are both individually open, but we are not mutually open. We are open, but about different projects. Indeed, it is appropriate to say that we are engaged in different projects all together. So again, it is not always the case that the voter fulfils this condition, again, to a sufficient extent to incriminate herself in the group project.

Lastly, do the voters share the relevant participatory intention? This depends on how we interpret it. Recall, the group of voters act jointly if they shared the fulfilment of this condition (along with the other conditions), if they fulfil a certain role in bringing about the outcome, where this role can be an instrumental role, and expressive role, or a normative role. First of all, as already shown in chapter 3, the instrumentality of the individual's act, is problematic to account for. It is not clear how the individual can make a causal contribution to the voting outcome. So it is difficult to see how the individual can fulfil an instrumental role. More plausibly, she has an expressive or a normative participatory intention, so we can presumably accept that this condition is usually fulfilled among some of the participants.¹⁰⁶ But like the other conditions, the fulfilment of this condition depends on how many share their particular understanding of the act they are engaged in. So it is not necessarily fulfilled by all the voters in favour.

Since strategic responsiveness is not fulfilled, and mutual openness are goal sharing, and participatory intention, are not necessarily fulfilled, there are reasons to believe that Kutz' account of joint action does not apply to the voting case. The main problem with the voting case is that it is to a large extent a private venture,

¹⁰⁶ Of course, since the participatory intention, as noted, incorporates the other three conditions in its description, the fulfilment of those conditions are also required for the fulfilment of the participatory intention condition. So this is assuming they have been fulfilled, which of course is not guaranteed.

which can make it difficult for the participants to fulfil even Kutz' fairly undemanding conditions for joint action. Many people presumably vote for their own private reasons (many people are e.g. reluctant to discuss who they vote for), at their own leisure (I do not coordinate my voting action), in a way disconnected from the actions of others, somewhat in the same way as when people play the national lottery. And even though they may vote for the same candidate, they are not necessarily doing it for the same reasons. This is not a criticism of Kutz' account aside from highlighting the limits of applying his account. It simply motivates examining his account to see if the conditions for his account can plausibly be loosened to potentially include individuals such as the voter.

7.4.2 Is joint action necessary for shared responsibility?

In this subsection I will show that we can plausibly loosen Kutz' conditions, without this affecting our intuitions of shared responsibility, at least in the cases I will examine. This is important, because as shown, his account will often be too demanding, at least with regards to the voting case and presumably similar cases. Particularly, I will show that it seemingly does not matter to my incrimination in some group-contingent harm, whether the *other* agents perceive me, or welcome me, as participating in their joint project in line with his conditions for complicity (particularly the first three), if I still do my unilateral part in the project. Note, that showing that we can loosen the conditions in certain cases, does not invalidate Kutz' account, nor any account of joint action with similar features. It does however help suggest that we can sometimes incriminate ourselves in a group project, even when we fail to fulfil the conditions for joint action proper. I will firstly compare two variations of the same case, and show that our intuitions of moral incrimination remain more or less unchanged across cases, even when one case is a clear example of Kutz' version of joint action, and the other is clearly not. This will indicate that something else than his account is incriminating us in the group project. After this, I will examine another example to show that the intuitions can generalize.

Firing squad – joint action, and joint responsibility:

I and nine other individuals have volunteered to participate in a firing squad (as shooters). We all know that the victim is innocent. Imagine that we are approximately equally skilled with our weapon. We all shoot with the intention of doing our best to kill the victim. We shoot, we all do various amounts of damage, none of us to an extent individually necessary or sufficient to kill the victim. E.g. imagine it took precisely five hits, no more no less, to kill the prisoner. Therefore, the combined effect was more than sufficient to kill the victim. Imagine further that we fulfil the conditions for having a participatory intention¹⁰⁷. Specifically, it seems plausible that me shooting my gun with the intention of killing someone, and actually hitting the target, is sufficient for having a participatory intention (it is perhaps instrumental and presumably expressive, and presumably also normative, i.e. in line with the norms of this particular military unit). Further, we all have the same goal (clearly). We further fulfil the rest of the conditions for joint action. I.e. we are strategically sensitive to the actions of the each other - e.g. we position ourselves in different places relative to each other. We shared the same goal – we at least all want to kill this prisoner, and this would indeed be satisfied if either one of us completes the deed. Lastly, we are mutually open, in the sense that each is able to reflect on the actions of others, and would accept their participation in a cooperative manner. I.e. they are all together in this project, and are acting in a cooperative spirit.

Whatever a joint action is, the above clearly fits our pre theoretical notion of a “joint action”. Even if the outcome in part due to my contribution, is causally overdetermined, insufficient, or even if my bullet happens to miss the target through some weird causal contingency, I imagine we share the intuition that all the shooters are morally responsible to the same extent for the death of the victim, and if we have any intuition that some are more or less accountable, this is only to a very small degree. This case seems like a perfect example where the

¹⁰⁷ I.e. we have an individual role, and a collective goal, where the role can be instrumental, expressive or normative.

fulfilment of the conditions for Kutz' complicity account explains our intuition of moral responsibility.

The following example is supposed to mirror the one above, with the exception that it is not in Kutz terms a joint action. After this I will examine another case which is quite different from the first, to show that there seems to be a pattern.

Firing squad – no joint action, but joint responsibility:

Imagine again a firing squad example, which is sufficiently similar to the above with some noted exceptions. In this case, I am aware of the other participants, but they are ignorant of me. E.g. imagine that there is a two-way mirror, so I can see them, but they cannot see me. However, I do not know they cannot see me. Further, imagine that they reasonably believe that they are shooting at a practice dummy, while I have a reasonable true belief that it is indeed an innocent prisoner. So *we* do not share a goal. *We* are not mutually open, because – let us assume – they would refrain from shooting if they knew it was a real person, and they would try to stop me best they could if they knew I was trying to kill the prisoner. *We* are presumably not strategically responsive either, because they would at least not adjust their behaviour in any way in furtherance of *my* goal. They would indeed try to block it by stopping me. Presumably *I* have a participatory intention in the sense of having an “intention to do my part of a collective act, where my part is defined as the task I ought to perform if *we* are to be successful in realizing a shared goal” (2000: 81). They however do not have the same participatory intention.

In this revised example, even if I reasonably believe, and it is true, that my shot will be insufficient to kill the victim when I pull the trigger, I think it is clear that our intuition of *my* accountability would be approximately the same as in the previous case, i.e. I would be morally responsible to the same extent for the innocent victim's death, and I would be deserving of the same resentment and presumably punishment. Regarding the other participants, their responsibility would depend on whether they were *culpably* ignorant of the fact that it was

indeed an innocent prisoner. If they are not, then they will go free. This suggests that whatever it is that is doing the work of incriminating me in the outcome, cannot be Kutz' account of joint action, or presumably *any* account of joint action. This is surprising, because cases like the above would be exactly the sort of case where the complicity principle and Kutz notion of joint action would be that which incriminates me, because it is a collective participation case (in the simple sense that it involves multiple individuals), and it is a case of overdetermination.

Of course, someone may object that it is mainly our direct contribution that is doing the work of incriminating the individual in the above case. E.g. they may suggest that the thing our intuitions are tracking, is the individual producing the outcome of *injuring* the prisoner, and not the *killing* of the prisoner. If that is the case, it is not my incrimination in the group project and the group-contingent outcome of the prisoner getting killed we consider me morally responsible for. It is instead the outcome of the prisoner getting severely injured. We can however control for that intuition by changing the example somewhat. E.g. we can imagine they are all shooting at a powder keg, standing next to the prisoner, where it takes five direct hits to blow it up. In this case, there is no direct injury which the individual causes directly. However, in that case our intuitions would still suggest that I am morally responsible for something just as severe as in the first case. But I have not injured the prisoner directly. I have at best "injured" the powder keg, which in itself is not all that morally relevant. This suggests that our intuitions are indeed tracking my responsibility for killing an innocent prisoner, even in the case where I fail to fulfil the conditions for joint action.

It is worth examining another example to see if our intuition is consistent across cases. I.e. our intuition that moral incrimination does not depend on joint action. I will construct an example where a group is engaged in a joint action, where you fail to fulfil the condition for joint action, but where you are still incriminated in the group act. I will merely construct one example here, i.e. the one where you *fail* to fulfil the conditions for joint action even though you are still incriminated in the act. It should however be easy enough for the reader to mentally fill in the

blanks. The moral stakes are also much lower, so if the reader has unclear intuitions, I suggest simply mentally increasing the severity of the outcome.

Rock Throwing:

You are walking up to a group of kids, and you see them engaged in, what to you seems like throwing rocks at a window. Filled with adolescent aggression, you decide to participate. Unbeknownst to you, they are actually not throwing rocks at the window at all. Rather, they are actually trying to hit a plastic Frisbee stuck underneath the tiles just above the window. Further, they are so gripped by their action, none of them see you *joining in*¹⁰⁸. But you do join in, and you apply as much effort as you can in your attempt at hitting the window. You also have a reasonable expectation that you (you all, the group) will be able to smash the window. You do not, however, hit the window at all, or do any amount of damage to the window or building. However, one of the boys is unlucky enough to do so (despite not trying to), and it breaks.

Now, you do not share the same goal, and they are not open to your goal of breaking the window. They also did not know you were there, and were not strategically responsive to you. So it fails to be a joint action on Kutz' account. You want to smash the window, so you participate in what you *think* is the group project, while they just want to dislodge the Frisbee, which is then *their* exclusive joint action. I imagine that we would consider all morally responsible for the outcome of the window breaking, even though *you* did not bring about the outcome. It was simply a careless act, which involved a lot of risk. You should all have known that it might end up smashing a window. Further, the owner of the window would be fully justified in requiring you all to pay an equal part of the destruction of the window, indicating retrospective indirect responsibility for its destruction (even if the owner came to know that you were *just* trying to break the window!). It seems whatever accountability we would plausibly assign to you, is no different from if the other children had been mutually open towards your

¹⁰⁸ Not meant to imply that it is thereby a joint action. 'Joining in' is just short hand for saying that you start throwing rocks.

participation, or had shared your goal. But you did not cause the destruction of the window, or even contribute to it in any causal sense. You missed the window, and it did not break because of something you did. If all the kids, including you, were appropriately performing a joint action on Kutz' account, then that could explain why you are incriminated in their act. But since it is not a joint action in that sense, something else must be grounding your inclusive responsibility here.

These examples give us further reasons to think that Kutz' conditions can be loosened and potentially ground responsibility in the voting case. Indeed, in the above cases, it seems that the pattern is, that something like the individual's exclusive¹⁰⁹ act plays the most significant part in justifying an ascription of responsibility. I.e. what seems to matter is that the individual conceives of *herself* as contributing to the project, regardless of whether there is any overlap between her and the other participants. The details of such an account needs to be developed in greater detail. I will sum up on this in chapter 8, when I construct an account of moral membership.

7.4.3 The problem with 'participatory intention'

I have so far shown that there are reasons to think that the conditions for Kutz' complicity account of shared responsibility can plausibly be loosened. Indeed, it seems that it is at times sufficient for my incrimination in a group project, that I have a belief that I participate in, or contribute to, the group act, when I act in accordance with this belief. What the relevant notion of "contribution" then is, is something I will discuss in the next chapter. In this subsection, I will show that one of the conditions for complicity is problematic, specifically the "participatory intention" condition. This is important, because if this condition specifies the conditions for when the individual performs a relevant action, and plays a relevant role in the group project, it has to do a good job at this. This is especially important if something resembling Kutz' account is going to be successfully applied to the case of political participation. Unfortunately this condition is

¹⁰⁹ Exclusive as opposed to inclusive, i.e. it depends on the individual's participatory intention regardless of whether anyone else also has this particular participatory intention.

problematic. The reason is that it is unclear how he can stipulate the conditions for a role in a project, which does not rely on the individual making a causal contribution to the project. If he cannot do this, this undermines his account, because his account is exactly a response to the problem of the individual not being able to make an individual difference to an outcome, and not being able to exercise control over that outcome. Further, it hinders its application in the voting case, because that case is exactly problematic because the individual is unable to make a difference to the voting outcome, or exercise control over whether it obtains.

The participatory intention condition is at the heart of the complicity principle. To restate:

“(Basis) I am accountable for what others do when I **intentionally participate** in the wrong they do or harm they cause. (Object) I am accountable for the harm or wrong we do together, independently of the actual difference I make” (2000: 122, emphasis mine).

It seems quite clear that this condition is supposed to be doing a lot of work in incriminating the individual in the group project.¹¹⁰ It specifies the particular contributory role the individual has in the group project. Strategic responsiveness, goal sharing and mutual openness are more akin to cognitive conditions, which when satisfied, secure that the relevant individuals in question have a sufficient overlap in their beliefs and attitudes towards each other, in order for what they participate in to be called a joint action proper. The participatory intention on the other hand, specifies the particular role each individual has to play in order to perform the actual contributory action in the relevant project, in the right way. This condition is extremely important, because without it, I could potentially be incriminated in a group project in which I am clearly not a part. I could e.g. be sitting on the other side of earth, talking with the airmen over Dresden, quite anachronistically, via Skype. In this case, I can still be incriminated in their joint

¹¹⁰ He does mean ‘participatory intention’, by “intentionally participate”.

action, even though my role fails to contribute to, or be constitutive of the harm they cause. Specifically, there could clearly be common knowledge or just mutual openness between myself and the other participants.¹¹¹ We could be said to be strategically responsive¹¹², if I time my call, and they time their response (or just switch on their screens) to my call in certain ways. Lastly, we can share a goal between us¹¹³, which would be satisfied if one of us completes it (even if I am unable to actually bomb anyone). However, none of this seems clearly sufficient to incriminate me in the relevant group project, certainly not to an extent anywhere near the pilots. If I am a member of the group, I am only an extremely passive member. Therefore he needs a further condition pertaining to me actually *doing something*, e.g. performing an action or making a contribution towards the project. This is why we need the participatory intention¹¹⁴ condition, because it specifies that I have to play a relevant role in the project.

Further, the participatory intention condition is also important, because the other conditions incorporate a reference to the participatory intention, i.e. the role or action of the individuals. Specifically, we have to be *strategically responsive* to the *actions* of other members. We have to be *open* to their *participation* in the project. We have to share a goal which is attained in light of the fulfilment of our

¹¹¹ Recall, *Mutual openness* requires that all members must have an attitude of openness to the fact that the other participants potentially reacts to their choices. We must both be open to the others' activity, and them to ours in the sense that we welcome each other's participation in the project, or at least would if we were aware of it.

¹¹² Recall, *strategic responsiveness* requires that all participants' intentions should be sensitive to the behaviour of the other participants', specifically their actions. This is required because "Jointly acting individuals do not merely act in parallel: each responds to what the others do and plan to do" (2000b: 5).

¹¹³ Recall, "two agents *share a goal* if there is at least one token activity or outcome, involving the actions of the other, whose performance or realization would satisfy the intentions of each" (2000b: 5, emphasis mine).

¹¹⁴ Recall, *participatory intention* is the self-regarding notion of acting with the intention of contributing to some groups act or outcome. By 'self-regarding', Kutz means that whether something is properly a joint action, is in part a matter of whether the individual herself intends for her own action to contribute to a collective outcome (2000: 74).

roles or the performance of our *actions*. So he needs some condition that specifies what the relevant action or role is, and that condition is the *participatory intention*. The centrality of this condition is best highlighted in the following quote: “the core, minimalist, notion of collective action as such requires **only** that individuals act on overlapping participatory intentions”. (2000b: 4, emphasis mine) In fact, Kutz goes as far as suggesting that the participatory intention may be the only condition necessary and sufficient for complicity, because he sees it as already encompassing relevant aspects of the other three conditions (e.g. 2000b section v). So it is absolutely central to his account.

Unfortunately, the ‘participatory intention’ is problematic. The problem is that it seems to rely on something Kutz explicitly rejects, namely causal responsibility. Recall, the original problem Kutz was trying to solve, was how to incriminate individuals in a group project, where their individual action fails to make a difference to that outcome and where they are unable to exercise control over the occurrence of that outcome. On his account, the solution to this problem was to incriminate the individual in a group action, where the group *did* make a difference to the outcome, and could exercise control over that outcome. If the individual is then incriminated in the group action, the individual then shares in the moral responsibility of the group, and the outcome the group brings about. This is what he means when he writes that “participation rather than causation is at the heart of both complicity and collective action” (2000: 134) But if that solution incorporates the problem the solution was trying to solve, it is clearly not a solution to the problem at all. It is not a solution to the individual failing to make a causal contribution, if it requires that the individual makes a causal contribution of some kind.

In order to show that the participatory intention condition either incorporates some notion of causation, or alternatively, is insufficient in specifying the relevant contributory role, recall that a participatory intention includes an individual role,

and a collective goal. The role can be expressive¹¹⁵, instrumental¹¹⁶ and/or normatively¹¹⁷ related to the joint action. I have already in the beginning of this chapter noted that the *instrumental* role seems to imply causal responsibility, so I will not go into much further detail regarding that interpretation. How about the expressive role? Does it incorporate causal responsibility? Well, either it does, or it is simply an insufficient account of the individual's contributory role. Whether it incorporates causal responsibility, depends on what Kutz means by 'expressive' here. As noted earlier, Kutz never elaborates the term, but simply writes: "[a] relation might be expressive if by doing one's part, one thereby exemplifies one's membership in a group or participation in an activity, as when by voting I express my membership in a political community" (2000b: 11). It seems he's relying on the notion of expressivity which is usually incorporated as an explanation for why voters vote, when the instrumental payoff is so small. Summed up nicely by Mackie, G.: "people vote not in order to pursue instrumentally the public good, but only to express inconsequentially a preference over an outcome, like cheering a sports team" (2011). If this is what Kutz means by expressive, this cannot suffice for incrimination in the group project. Simply expressing one's sympathy with an action or outcome is far from sufficient to incriminate one in that action or outcome.¹¹⁸ He could of course mean something like expressing a commitment to do one's part in the project. But that just raises the question of how one would do one's part. The most obvious way of doing one's part is by making a causal contribution to the project in some way, so that would not solve the problem.

¹¹⁵ Regarding the expressive role, he writes: "The relation might be expressive if by doing one's part, one thereby exemplifies one's membership in a group or participation in an activity, as when by voting I express my membership in a political community" (2000: 82).

¹¹⁶ Regarding the instrumental role, he writes: Contributory relations might take instrumental form if what the agent does helps cause the collective outcome [...], or if the agent's part is a constitutive element of the groups act" (2000: 82).

¹¹⁷ Regarding the normative role, he writes: "the relation might be normative if one performs one's part because of norms internal to some group or institution that demand certain behaviour (I wear a dark suit as an IBM employee)" (2000:82).

¹¹⁸ It may of course shed a negative light on the person in question, and it may lend credence to the thought that this person *would* contribute to the outcome, if the relevant situation obtained.

Perhaps it is not the expression of a commitment which is relevant, rather it is *making* the commitment itself. In that case, I have made a commitment to do my part. But again, “doing my part” refers to an action or a contribution of some kind, so he still needs to give an account of action or contribution, which does not rely on making a causal contribution. He could also mean to allude to the last line in Mackie’s quote, of cheering for a sports team. Cheering for a sports team may presumably help the players perform better. So if I am expressing my support for a given project by cheering for the other participants, then I am presumably helping them in a way which may incriminate me in that project. However, this is again a way of my having a causal role to play. I am affecting their behaviour by my action. So this also reintroduces the notion of causal responsibility, or in Kutz’ terms, making an individual difference. Alternatively, by ‘expressive’, Kutz could mean performing actions which are expressive, as something different from simply expressing an opinion or a commitment. E.g. me shooting my gun at someone, but missing the target, may certainly be expressive of my intention of killing that person, or as it were, expressive of my role in bringing about his death. But in this case there is an actual action, and I presume it is that action (the, intentional trying to accomplish something and acting on it), rather than the mere expression, which incriminates me somehow. He could also argue that by expressing something, that makes it appropriate to consider the agent’s behaviour predictive of future behaviour, or predictive of an intent to produce harm. But then that expressive role is still piggybacking on some notion of causal responsibility, because it is the instrumental contribution which is being predicted, and the expression is irrelevant in itself. All in all, it is difficult to see what work the expressive understanding of role in participatory intention is doing here, and why having such a role would incriminate someone in a collective harm, even when we also fulfil the rest of his conditions for joint action.¹¹⁹

Lastly, Kutz also notes that the participatory intention can be fulfilled normatively, if our role in the project is normatively defined somehow. Unfortunately, Kutz does not give a substantial account of how the role in participatory intention

¹¹⁹ Mutual openness, strategic responsiveness, goal sharing.

could be filled normatively, but the example he gives seems to imply that he means acting in accordance with something like rules of conduct. I then fulfil the participatory intention condition for moral incrimination if my behaviour is in line with the particular norms of e.g. the military who bomb Dresden. But this seems problematic also. I can e.g. be at a military base in a different country and fulfil the relevant military norms, without this clearly incriminating me in the particular joint venture in any clear sense at least. Of course acting in accordance with certain norms may be indicative of a tendency to act in certain ways. Someone who accurately fulfils the role of being a soldier, may only be doing so, if she will be willing to discharge her weapon, should certain circumstances arise - even if she is not at a particular moment engaged in combat. E.g. if she is put in a combat situation, and runs away, she is not fulfilling that role. But again, here it seems the significance of the normative role, is again explained in terms of a *potential* or *predictive* causal contribution. It is because I have a role where I may at some point perform an action directed at the outcome that I have incriminated myself. So it seems again that Kutz is embedding some notion of at least potential causal contribution in the participatory intention condition. Admittedly, the normative interpretation of role may be sufficiently open to interpretation to actually incorporate some notion of role, which is not actually, potentially, or predictively *causal*. It is however not clear to me how.

It is possible that there is some interpretation of the participatory intention condition which does not rely on a causal contribution in any sense. But it is unclear at this point what it is. As it stands now, it is not clear how he can avoid committing himself to the requirement that the individual makes some causal difference to the project she engages in somehow. Indeed, it seems that some notion of making a causal difference is needed for an account of the role the individual has to play in a project, in order to be incriminated in that project. This is a problem for his account, since it is explicitly a response to the problem of accounting for moral incrimination in cases such as the overdetermination case, where our causal contribution is difficult to account for. Therefore – if making a causal difference is off the table – we may have to reject his account, since it does not give a viable alternative account of the participatory intention. Or, we may

accept that making a causal contribution, at least potentially, is a necessary condition for joint action, and shared responsibility as complicity. That participatory intention seems to incorporate causal contribution, is also problematic for applying his account to the voting case, because the voting case is exactly a problem case since none of the voters seem to make a causal contribution to the relevant outcome, aside from in the most unlikely of circumstances. Therefore, if something like Kutz' account is going to be applied to the voting case, it is not sufficient to loosen the other joint action conditions. It is also important to give an alternative account of how the individual should contribute to the voting outcome, because as argued in chapter 3, having to make direct causal contribution is off the table.

7.5 Complicity and gradability

Another reason why Kutz' notion of participatory intention as it stands is unintuitive, is that it can be too inclusive. Someone can be incriminated in a group project, even when they clearly should not be, or alternatively, where Kutz' account is unable to explain the extreme variation in our intuitions regarding their moral evaluation, e.g. how blameworthy they are or what punishment we would consider them deserving of.

Recall, Kutz denies that the individual difference principle should play a role in accounting for the participants' accountability, because "we will quickly fall into contradiction, since the basis for complicitous accountability is inconsistent with the individual difference principle" (2000:147) a principle he explicitly rejects, because he wants to show how we can be responsible "independently of the actual difference [we] make" (2000: 122) , again highlighting that the individual difference we make is unavailable to draw on in his account in general. This means that that difference is unable to figure in an account of our varying intuitions of accountability between different agents who all are incriminated in the same outcome. E.g. if causal contribution was a condition for moral incrimination, we might be able to say that those who made the most significant contribution, ought to deserve more of the blame and harsher punishment.

The following example shows that Kutz' account seems to be too inclusive *as is*, i.e. it lets us hold someone morally responsible for the relevant outcome, where they either clearly are not, or where his account at least fails to accurately account for our intuitions of gradability:

Ideologically inclined cleaning lady:

Say I am an ideologically inclined cleaning lady, working at the administration building of a military general who gave the order to attack Dresden. I have the relevant participatory intention of bombing Dresden, because I took this job to make a contribution to the war effort (because I simply hate all Germans), and my job plays a role in furthering the war effort, because by cleaning the office, this lets the general focus on his job of drawing out effective plans etc. Whatever Kutz' interpretation of a *participatory intention* we accept, it will presumably be fulfilled. E.g. we can imagine it is fulfilled instrumentally, where my role is defined by my instrumental act of cleaning. We can also imagine it is filled expressively - this is my way of expressing my support. Now, am I accountable? I have the participatory intention, I am both expressively and even instrumentally related to the outcome. We are all aware of, or *open* towards each other's participation. We *share certain goals* (and would accept its completion by any participant) where keeping the general efficient at his job would be such a goal. Lastly, we are strategically responsive (I e.g. adjust my cleaning schedule to the general's office hours, and he puts his chair up, so I can clean underneath it).

Though I am included as part of the joint action, I imagine that whatever our intuitions of reactive attitudes, and punishment, they will be entirely different from those directed at the bomber pilot who dropped his payload on Dresden. I imagine our intuition here suggests that the range is simply extreme. Whatever punishment or reactive attitudes we consider me deserving of, is at another end of the spectrum from the bomber pilot or the general, and to claim that they are all equally responsible for the outcome will seem odd. Presumably we will not

consider the cleaning lady liable to anything but shaking our head at her odd disposition.

Now, Kutz does not deny that the causal difference we make, can also play a role in an ascription of supplementary direct moral responsibility. I can be responsible both for my complicity, and independently on a direct account. But my complicity is not grounded in my causal contribution. Many ordinary cases are not cases of complicity with individuals participating in group action, but are instead simply cases of some individual intentionally directly producing or trying to produce some outcome. In that case, accountability may be appropriate with regards to the actual difference the agent makes directly. So if we ignore Kutz' claim that the bombers fail to make a difference to the outcome (due to the destruction being overdetermined), and assert that part of the difference in our evaluation of the agents in question comes down to these pilots individually producing a large amount of direct harm, we should still hold that part of our moral evaluation comes down to what is conferred on the participants in terms of shared responsibility as complicity. If not, what is their shared responsibility really accounting for? But if all the resentment we can then muster towards the cleaning lady amounts to mild ridicule, that must imply that complicity itself – implausibly – does not account for any substantial degree of moral incrimination in the outcome. This shows that his account is either implausible, or it shows that complicity, only grounds an insignificant moral evaluation on the agent, even in an outcome as severe as this (thousands of innocent people are unjustly killed). Alternatively, he needs to account for why we have varying intuitions of accountability between complicit individuals, which is not explained by their causal contribution.

Kutz may be able to account for the psychological aspect of why there may be *some* difference between our evaluation of various agents, but assuming the participants are not coerced into participating, or the cost of defection is at least very minimal (assume that's the case for the cleaning lady – she can just quit), we should assume that her moral responsibility for the whole outcome would at least play a significant part in the moral evaluation, which it does not seem to do. And

we lack some intrinsic element of his complicity account which can account for our varying intuitions. This again suggests that the cleaning lady is either not incriminated in the project, despite fulfilling Kutz conditions, or, it suggests that his account is unable to clearly account for gradability in our moral evaluation.

Saba Bazargan has given a supplementary account of how we can interpret participatory intention to solve the problems of gradability on Kutz' account (Bazargan, S. 2013). Bazargan argues that the *role* people fulfil in the project serves as a basis for gradability of moral responsibility (Kutz himself discusses this, but Bazargan argues the point more clearly). Bazargan argues that one's moral evaluation depends on the "type" of role one has in the project, where the more central one's role is with regards to the project, the more complicit one is, and the more severe a moral judgment one is deserving of. "All things being equal, the greater the degree to which one is supposed to contribute to a cooperative act, the more prominent one's role in that collective act" (ibid: 189).

Bazargan gives an example of different roles in a military rescue operation to illustrate the gradability via roles: "ineffective soldiers in the squad rescuing the POW will bear greater inclusive authorship for the rescue of the POW than will, for example, sailors aboard a [nearby] minesweeping vessel – even if their respective contributions to the rescue of the POW are on a par. The soldiers in the squad bear greater inclusive authorship since they have roles that feature more prominently vis-a-vis the end of rescuing the POW" (ibid: 189). Though this seems promising, it is not at all obvious how he can account for 'prominence'. Bazargan tries to clarify this by stating that one's role "is designed to contribute to a degree far greater than [the others]"(ibid: 189), but this of course just pushes the question back, so we end up asking what "design" refers to here. Bazargan also, just like Kutz, rules out causation as an indication of *role design*, for good reason, because the problem is still that individuals who fail to make a causal contribution individually, can be responsible for a collective harm (ibid: 188). So this is not a good solution to the problem of gradability, because it seems to be question begging. We surely do have intuitions of how different people incriminate themselves differently, e.g. the cleaning lady was intuitively far removed from the

outcome compared to the airmen, but in order for this account to be applicable in the voting case, we need a clear criterion which will determine why we have different intuitions in these cases. And central to this project, we need to account for our varying intuitions in small vs. large scale voting cases. In the next chapter I will produce an account which does explain our intuitions in such cases.

7.6 Summing up this chapter

Kutz' account was an attempt to specify how the individual could incriminate herself in a collective harm, in cases where she fails to make a difference to that harm and fails to have control over its occurrence. His suggestion was that when a group of people fulfil the conditions for joint action, they can then be morally responsible for the outcome of that joint action, again, even if they fail to make a direct contribution to that outcome. Though seemingly promising, is not clearly applicable to the voting case. The reason is that it is not clear that the voters will fulfil the conditions for joint action. However, I have argued that we can plausibly loosen the conditions, and still incriminate the individual in the relevant project, which, as I will show in the next chapter, will allow us to attribute moral responsibility for policy outcomes to the individual. It seems that something like making a unilateral contribution to the outcome, is sufficient for moral incrimination, i.e. I do not have to fulfil the conditions for joint action proper to be incriminated in a project. Unfortunately, Kutz' condition for when the individual can be said to be making a contribution was also problematic, and thus not readily applicable to the voting case. Specifically, the *participatory intention* condition, which specified the individual's contribution, seems to include something like a requirement that the individual makes a difference to the project she is incriminated in. It does this because it specifies the relevant role the individual should perform, but that role was difficult to specify without it resting on the individual making an actual causal contribution. I.e. it contradicted the idea that an individual can be incriminated in a project, even if she fails to make a difference to that project, or fails to exercise control over its occurrence. This is a problem, because Kutz' account is precisely a response to the problem that the individual fails to make a difference in certain collective action cases. So having this as a requirement for incrimination in the group project is a problem, since it

reintroduces the problem it is a response to, i.e. the problem that the individual fails to make a difference or fails to exercise control in certain cases. Further, it is again problematic with regards to applying Kutz' account to the voting case, because in that case, as argued in chapters 3 and 4, it is not clear that the voter actually makes a difference to the outcome. So a successful account for the sake of the overall project, needs to either explain how the voter makes a difference in the voting case, or stipulate an alternative to Kutz' participatory intention condition for moral incrimination. As noted, this condition is extremely important, because it specifies how the individual has to actually participate in, or contribute to the project she is incriminated in. So a successful account presumably needs to give an alternative account of how the individual can contribute to the voting project. Lastly, another problem for Kutz' account is that it is either too inclusive, or it lacks a clear criterion for gradability between participating agents in collective projects. The cleaning lady was either implausibly incriminated in the group project, or alternatively, Kutz' account fails to give an explanation for why our intuitions of accountability vary so much between participants in a given case. So a successful account of shared responsibility needs to either be able to exclude someone like the ideologically inclined cleaning lady, or explain our intuitions of accountability gradability in cases such as that.

In the next chapter, I will construct an account of shared responsibility, which lets us incriminate the individual, even if she fails to make a direct contribution to the outcome.

8 Moral membership

8.1 Introduction

In this chapter, I will build on the preceding chapters, and construct an account of indirect moral responsibility which can be successfully applied to the voting scenario. I will denote this account a 'moral membership' account. This label highlights that this particular account is somewhat distinct from those examined so far. Further, the label is meant to signify that it is the *membership* of a

particular “project”, which grounds the individual’s moral responsibility, and not e.g. that one is engaged in a joint action proper.

The account I will propose is very reminiscent of Kutz’ complicity account, in that it incorporates a specific participatory aspect of his account. Specifically, it incorporates the necessary condition that *I* participate in what I foresee as being a collective project. It is however distinct from his account, because it allows us to be indirectly incriminated in a project that brings about a harmful outcome, even if we fail to fulfil the conditions for joint action proper. In the previous chapter, I argued that the conditions for joint action and shared responsibility on Kutz’ account, could plausibly be loosened while maintaining the moral incrimination of the individual. This was important, because the voter did not clearly fulfil the conditions for joint action. On Kutz’ account, a number of individuals (very roughly) fulfil the conditions for joint action, and shared responsibility for an outcome, when there is an overlap between the participants’ intentions to bring about the outcome, welcoming attitudes towards the participation of one another, and a fulfilment of a contributory role directed at the outcome. However, as I highlighted, there are cases where it seems sufficient for *my* incrimination in a harmful outcome (again, very roughly), that *I* have the intention to produce the outcome, and *I* fulfil some contributory role in bringing about the outcome, regardless of whether there is complete overlap between me and the other participants and regardless of whether I am directly causally efficacious. Indeed, if e.g. I am aware that the project we are engaged in, will result in an enormous harm, but everyone else but me is unaware of this, I can still be responsible for the harm, even when the other participants are exculpated for their contribution (e.g. because they were non-culpably ignorant that a harm would obtain). Further, it seems I can even be incriminated in a project, when the participants are not open towards my participation, e.g. where they have clear adversarial attitudes towards me to the extent where they would even try to sabotage my participation if they were able to. This suggests that it is – at least in some cases of shared responsibility - not the joint action, but instead the individual’s participation in, and her conception of, the particular project she

actually participates in, which grounds her morally responsibility. I.e. it is her *membership* of the project.

The most important development compared to the accounts examined in chapters 6 and 7, is that I invoke a novel notion of moral-responsibility-grounding indirect contribution. When the individual makes such a contribution to a particular project, she is a moral member of that project, assuming she also fulfils the cognitive condition introduced in chapter 2 and discussed in chapter 4. I.e., if she is aware of her contribution, and the significance of the project, or ought to have been aware with regards to both. Specifically, an individual contributes to a project if she is likely to be an interdependent part of the set of causes which actually bring about the outcome of that project. Not because she directly affects the outcome (in many cases she does not), but because she helps establish and/or sustain the project which causes the outcome. It is not applicable to all cases of shared responsibility, but it is highly suitable in the voting case, and cases like it. I will flesh this out in more detail in the rest of this chapter, but here simply give a revised example from earlier to give an idea of this account and this notion of contribution. After this, I will lay out the overall structure of the chapter.

The example is under-described, and not at all cleaned from conflicting intuitions, but it should give an idea of the moral membership account applied.

Bridge collapse 2:

Imagine you drive up to a bridge which is closed midway across, due to heavy winds. The bridge has many cars parked, unable to turn back, waiting for the bridge to be re-opened. It so happens, that you are a bridge engineer with a death wish. You recognize that the bridge is moments away from collapse due to a fragility from an otherwise hidden structural deficiency, and the heavy winds. You decide to drive onto it, knowing that the combined weight of all (say) 50 cars and trucks, along with a strong gust of wind, is more than sufficient to make it collapse any moment now. As you predicted, the bridge collapses, killing most of the drivers and passengers, including yourself.

Imagine that it took precisely 40 cars and trucks to reach the bridge's collapse threshold when the wind hit. In this case, you are a moral member of the *project* of the bridge's destruction and killing everyone (presumably the only member), because you were a likely part of the causally efficacious set that brought about the outcome. The causally efficacious set was the exact number of cars required for the bridge's destruction. We do not know whether *your* car was part of the actual threshold, but it is still appropriate to claim that your car helped sustain the project because it was at least *likely* that your car was part of the threshold (80% chance). Since you know this, and since it is true that you were indeed a likely part, you are a moral member of the project. Because you are a moral member, you are morally responsible for the outcome of the project, even if you are the only moral member. You are presumably the only moral member, because only you fulfilled the cognitive condition for moral responsibility. This example is distinguishable from cases of moral incrimination by way of Kutz-style complicity because it is not a case of joint action, since the other contributors (drivers) are unaware, and let us assume, extremely hostile towards the bridge collapsing, and your contribution to that end (because it will kill them). I will develop and defend the moral membership account throughout this chapter.

Regarding the overall structure of this large chapter: In section 8.2 I will sum up four of the most central observations from this project so far, which will be incorporated into an account of moral membership. Doing this will be helpful in establishing the conditions for the account. Specifically, 1) an account of shared responsibility as moral membership should not require that the individual makes a significant direct contribution to the relevant outcome itself. This is a natural response to the failure of direct accounts to account for moral responsibility in certain cases of collective participation, e.g. the voting case. 2) Shared responsibility as moral membership does not require 'participatory reciprocity' among the participants. This is a rejection of the notion that *my* moral incrimination in a given project, requires that the other participants are aware of me, or that they are open towards my participation. Further, it is the rejection of the notion that we ought to have the same conception of the project we are

engaged in. I will elaborate in detail, but this is the lesson from the previous chapter. In some cases of shared responsibility, such as the firing-squad case, I can be morally responsible for the outcome of the innocent prisoner dying, even if I am the only one who is aware that this is in fact a human being, and even if the other participants would be opposed to the killing of this person, if they knew it was indeed a human being. They are not open towards my participation in the project according to my conception of it. But I am still morally responsible for the outcome of the project nonetheless. 3) An account of moral membership ought to be able to account for our intuitions of gradability. Specifically, in various comparable cases, we will have different intuitions regarding the participant's accountability, e.g. their blameworthiness and punishment. An account of shared moral responsibility as moral membership should be able to explain at least part of this variation, at least in some cases. 4) An account of shared moral responsibility as moral membership requires that the individual makes some contributory action. I have highlighted this point with regards to all indirect accounts of moral responsibility examine in chapters 6 and 7, but it seems to be a common feature, that for someone to be incriminated in a collective project, this requires that the individual participant contributes to this project in some sense, even if not a direct contribution to the outcome of that project.

In the large section 8.3, I will propose a way in which the individual can become a moral member through performing a specific contributory action. As I will argue, (and which has been implicit or explicit in the previous accounts of indirect moral responsibility), in order for an individual to incriminate herself in a group project, she has to perform *some* indirect contributory action. There are possibly various ways in which the individual can perform such an action, but I will focus on one that is particularly relevant to the voting case. I will arrive at this by examining an account by Seamus Miller, who argues that we can incriminate ourselves in a group project, if a certain type of *interdependence* obtains between the participants which sustain the project. Though his account is not directly applicable to the voting case, I will build on it, to construct what I refer to as *threshold interdependence*. When threshold interdependence obtains between a group of people, they can then be said to make a contribution to the relevant

project which produces the relevant outcome, even when they fail to make a direct causal contribution to the outcome itself, e.g. even when the outcome does not depend on them making an individual difference to it. If they make such a contribution, they then qualify for moral membership, assuming they fulfil the cognitive condition (e.g. are aware of their contribution or ought to be aware of it). Threshold interdependence builds on work done by Richard Tuck (2008), who argues that ‘contribution’ in a threshold case, such as the voting case, can be understood in a special way. Particularly, we contribute to the group act, when it is *likely* that we were part of the causally efficacious set that brought about the outcome. From this I will argue, that we contribute to a given project that causes the relevant outcome, if it was likely that our contribution was part of the causally efficacious set that brought about the outcome. When this happens, it is appropriate to say that there obtains an interdependence among the participants, where the individual participants sustain the project that brings about the outcome. By sustaining it, they can then be morally responsible for the outcome of that project. ‘Contribution’ does not here depend on us making a direct causal contribution to the outcome itself, but rather, it depends on us making a contribution to the project, which in turn brings about the outcome.

In section 8.4 I will put everything together to construct the full account of moral membership. In the last chapter (9) I will apply it to the voting case.

8.2 Observations from the previous chapters

I have reviewed various accounts of moral responsibility with two aims in mind. 1) Ascribing moral responsibility to the individual citizen through her political participation, specifically her voting action. 2) Examine what is required for an account to successfully do so. So far I have failed in 1). In this relatively short section, I will highlight some important observations made throughout this project, in order to shed light on 2). The goal of doing so is to inform the construction of the moral membership account, which can be successful in incriminating the individual in the voting case. Highlighting these observations will also be helpful as a way of taking stock of the some key points in the discussion so far. After having highlighted these observations, I will construct the account based

on these. This section includes four observations. I will highlight them in turn. The last observation is most central to establishing an account of moral membership, so I will engage with it to a more significant extent than the other three.

The **first observation** is that a successful account of moral membership should not require that the individual makes a significant direct contribution to the relevant outcome itself. This is the obvious conclusion derived from the discussion of direct moral responsibility, and the limit of such an account with regards to large scale threshold cases. The problem is the individual fails to make a significant direct contribution to the relevant outcome in such cases. In cases of direct moral responsibility, the individual is morally responsible for an outcome, in part because she is causally responsible for it. She is causally responsible for it, because the outcome depends on her action (or plausibly omission). In cases such as the large scale voting case, the outcome does not depend on the contribution of any individual, and no individual makes a significant direct contribution to the outcome, except in extremely unlikely scenarios. Because of the difficulties in attributing the voting outcome and policy outcome to the individual voter on a direct account, I instead examined indirect accounts. These accounts were specifically defined as accounts of moral responsibility which did not require that the individual, as an individual, made a significant direct contribution to the outcome. However, though none of them has been successful so far, the conclusion drawn from examining direct accounts still remain: a successful account should not rely on the individual making a significant contribution to the outcome itself.

The **second observation** is that I can be a moral member of a project, even if there is no 'participatory reciprocity' between myself and the other participants. What I mean by 'participatory reciprocity' is at least 1) that the individuals involved in the relevant project are aware of each other, or that they at least are open towards the potential participation of each other. 2) That they have approximately the same conception of the project they are engaged in. This observation is supported by the fact that our intuitions of moral responsibility seem quite unaffected when we remove participatory reciprocity from the

participants in joint action cases of clear shared responsibility. This was most clearly illustrated by the firing squad case which is comparable to the voting case.¹²⁰ Here it seemed sufficient for *my* moral incrimination, that *I* had a particular true conception of the project, and that I was open towards the inclusion of the others in bringing about the project. However, the fact that these features were not reciprocal, did not affect my moral incrimination in the project, as *I* reasonably conceived it, even though it undermined the project as a joint action. I will show throughout this chapter that this intuition does indeed generalize. I will assume that participatory reciprocity is not a necessary condition for moral membership. A related question is of course, what is a “project” then if it is not a joint action, and if it does not involve participatory reciprocity? The answer is that it is the observable leftover structural relationship between the individuals, when we take away their participatory reciprocity, i.e. when we take away their strategic responsiveness, goal sharing, mutual openness, and their *shared* participatory intention (to e.g. use Kutz’ criteria). When we remove these from the joint action, then there will still be a multitude of individuals which together in some conjunction, as a set/group causally produce the morally relevant outcome. Of course, it is an interesting discussion how we should individuate such projects, but safe to say there will be obvious cases such as those I have examined previously. In those cases, this conjunction of individuals can still be understood as forming a *project*. If the individual understands (or ought to understand) that she is engaging in such a project, she is – as I will argue – open to an ascription of moral responsibility for the outcome of that project (e.g. as is the case of the shooter in the firing squad behind the mirror from the last chapter).

Returning to the voting case again, we can see why the denial of participatory reciprocity is important for moral incrimination in that case. In the voting case, we can imagine a scenario where I vote for candidate C with the aim of bringing about outcome O (a foreseeably harmful outcome) as a consequence of project P, but where others vote for candidate C with the aim of bringing about outcome $\neg O$ (a foreseeably benign outcome) as a consequence of project $\neg P$. Therefore, we do

¹²⁰ If it is construed as a threshold case, where no individual can bring about the outcome herself, and the sheer number of participants overdetermine the outcome.

not have the same conception of the project we are engaged in.¹²¹ Let us assume then that I am right about C ultimately leading to O, and they are wrong about C ultimately leading to $\neg O$. If they knew I was right, they would presumably be prepared to sabotage my contributory action if they were able to (at least by voting against O). If participatory reciprocity is a necessary condition for moral membership, then I am not a moral member of the project that brings about O, because I could not have brought it about by myself. Presumably there will usually be others who have the same beliefs as me. So I may be grouped with all who shared the same conception of the project as myself. However, we can still imagine a case where we are a significant minority, and would not be able to bring about the outcome alone. Therefore, an account of moral membership will be more readily applicable if it does not rely on participatory reciprocity, since this allows for broader inclusion in morally relevant projects, and thus the outcomes thereof. And indeed, it seems to be the case that the individual in some cases can share in the responsibility of a collective project, even if there is no 'participatory reciprocity' between her, and the other participants and where she could not have brought about the project and the outcome by herself.

The **third observation** is that an account of moral membership ought to be able to account for our intuitions of varying degrees of accountability in relevant cases. As noted, in cases of collective participation and also documented in the chapter on causal responsibility, our intuitions regarding accountability seems to vary with the number of participants, or with the particular role the individual plays in the collective project (e.g. recall the culpable cleaning lady from chapter 7 (section 7.5). This seems immediately clear in some of the voting cases examined thus far. If it is a small group voting for or against some large harm, our intuitions will presumably be that whatever reactive attitude or punishment they deserve for helping bringing about this outcome, will be significantly more severe, compared to a case where millions of people vote for or against the same harm. Obviously the voting case is not a good example with which to demonstrate this issue,

¹²¹ We are aware that we are voting, but voting is merely the tool here, and not the conception of the project we have. The project we are engaged in and have a conception of, is the project of bringing about some ultimate outcome, i.e. a policy outcome.

because it is the particular case that is under scrutiny, where we lack clear intuitions going in. An account of moral membership is supposed to be defended independently and then applied to the voting case in order to let us trace out the relevant verdict regarding the moral responsibility of the voters. But the same carries in other cases. E.g., imagine the comparable case of overfishing. In one case, we have 10 highly efficient, quite automated boats, each controlled by their individual owner, which together, due to their combined effort, produce the outcome of destroying the the local ecosystem. Assume then that this outcome can only be brought about by a group effort (e.g. it requires at least 7 boats). Assume this is a morally relevant outcome, e.g. the local inhabitants starve because of this. Glossing over certain details, we will presumably have a reasonably strong intuition that they are all accountable for the outcome to a severe extent. Then compare this to a scenario which is comparable with the first, with the exception that the 10 boats are not as automated. Instead, each boat takes specifically 100 fishermen to function to produce the same combined result. So each boat is indeed managed by 100 fishermen, who also happen to be the collective owners of their individual boat. In this case we now have 1,000 individual participants instead of 10, and presumably our intuitions regarding their culpability will vary across cases even though the only relevant difference seems to be the number of participants. In the latter case, many will presumably share the intuition that each individual is less blameworthy. Or if this is unclear, then it will presumably be clearer if we keep adding even more fishermen to the boats. Part of this is of course due to diverging liability across cases. In one case, the cost is split between 10, and in the other between 1,000, while the actual damages are constant across cases. But presumably, we will also have intuitions of varying degrees of blameworthiness. An account of shared responsibility as moral membership ought to be able to explain this variation. The challenge is accounting for gradability, when gradability is not tied to the individual's direct causal contribution. In two comparable cases where we have varying intuitions regarding the accountability of the individual, it is often straightforward to explain this variation in terms of the variation in the foreseeable and actual direct contribution the individual makes. But in cases like the voting case, where we assume the individual fails to make a direct contribution to the outcome, such a

contribution will not provide any insight into the individual's accountability. So a challenge for an indirect account is to explain why our intuitions of the participants' accountability will vary from cases to case.

The **fourth observation** is that shared responsibility as moral membership requires that the individual performs some contributory qualifying action. What is meant by 'contributory qualifying action'¹²², is some action which can appropriately be understood as contributing to one's conception of the relevant project, which qualifies the individual for *membership* of that project. Membership which then in turn may justify an ascription of moral responsibility for the outcome of that project. In the subsequent section I will discuss the particular notion I consider necessary for moral membership, and here simply document that it seems to be a feature, or a missing feature of the indirect accounts examined thus far.

It is of course difficult to substantiate this assertion, but none of the proposals examined so far were able to account for an ascription of shared responsibility without some approximation of it. With regards to Sverdlik's account, we indeed lacked a clear account of how the individual ought to contribute to the outcome which was incorporated into the agent's intention, in order to be incriminated in that outcome. Sverdlik suggested that it was necessary that the outcome, which is incorporated into one's intention, comes about as a result of one's action. However, he never specified how this outcome would come about in cases of collective participation. With regards to Abdel-Nour's account, he did incorporate a notion of contribution. Specifically, on his account we are responsible for an outcome, if there has been established a continuity between us and the historical nation, which is exemplified by our pride in the accomplishments of the nation. When this continuity has then been established, we *are* the nation, and what we then contribute to, is what the nation contributes to, or has contributed to. As argued however, it was doubtful whether the suggested identificational link was

¹²² The term "qualifying action" is here borrowed from Mellema, G. (2006), who uses it in the same way, to refer to actions which precisely qualify the individual participant for membership of some harm.

sufficiently strong to support the conclusion that the individual citizen brings about the relevant outcome, to an extent which makes us morally responsible for that outcome. Pasternak's account did not explicitly require that the individual contributed to the outcome. Instead she stipulated that the individual ought to be a member of the state according to its own rules, and that the individual ought to be aware of her membership of the state. Pasternak's account was not an attempt to ground moral responsibility for the outcomes of the state, but rather to justify how the state can pass on its liabilities to the citizens. However, I argued that something more than simply being a member according to the states formal rules (and being aware of this) was needed for moral responsibility. Surely it is plausible that I have to be aware of my membership in a project which produces an outcome, in order to be incriminated in that outcome, but *membership* presumably needs to be grounded in something else than a formal set of rules. I suggested that what was missing from Pasternak's account, was that the individual performed some contributory action, which would then ground her *moral* membership, regardless of whether she was *formally* a member of the particular project or institution. Lastly, Kutz also included the requirement that the individual had some contributory role to play in the joint action, in order to share in the responsibility for the relevant outcome. Indeed, Kutz seemed to implicitly be relying on some notion of *direct* contribution as a necessary condition for moral incrimination. However, though I argued that this was problematic, he would presumably be willing to accept that some other notion of contribution would satisfy the individual's role. With the above in mind, it seems that a central feature to an ascription of indirect - and of course direct - moral responsibility, requires that the individual performs some contributory act of incrimination.

One seeming exception to the observation that shared responsibility requires contribution, is the not-actual-contribution-but-still-playing-an-important-role idea of participation. So I will examine this in detail, and again show that it is difficult to understand how the individual can incriminate herself in a project, unless she makes some contribution to the project, either actually, potentially or reasonably predictively. Because we seem to be forced to accept some notion of

contribution in this case, that adds credence to the observation that moral incrimination in a group project requires some notion of contribution to that project. The best example of such a view I am aware of, comes from Saba Bazargan (2013) whom I mentioned in the previous chapter. Bazargan tries to argue that the individual can be incriminated in a group project, merely by having a non-contributory *role* in the project. He gives the example of a bank-heist where a lookout has a role to play, fails to make a clear contribution, but intuitively still is culpable simply because of fulfilling the relevant non-contributory role. I will simply quote his case in order to discuss it:

Heist:

“a criminal mastermind puts together a plan for robbing a bank. She hires five individuals, each of whom agrees to participate in the robbery. The recruits are made aware that part of the plan is to kill the witnesses in the bank. The mastermind does not physically participate in the robbery - instead, she provides the plan, the layout of the bank, the equipment, etc. One of the recruits, J, is stationed on a second floor balcony above the bank, as a look-out. Her role is the least important. The mastermind would have commenced with the plan even without a look-out. Suppose that J is not a very effective look-out—in fact, she falls asleep on the job. Fortunately for the robbers, J’s incompetence has no negative effect on the robbery, though her participation does not causally contribute to the robbery or murders either. The plan succeeds, and two witnesses are killed [...] though J causally contributed nothing, she bears some liability for the murder of the witnesses and the theft of the money.” (Bazargan, S. 2013: 182-3)

I will assume that most people will agree that J shares in the moral responsibility here. Bazargan tests the strength of this intuition by asking whether it would be ok for the police to shoot and potentially kill J, if her death would result in other innocent lives, e.g. those of the bank-heist hostages, being saved.¹²³ If we assume

¹²³ His example is that they shoot him, and he falls from the balcony which then startles the robbers, and makes them give up their robbery.

this is the only way to spare the lives of the innocent hostages, we would probably find this acceptable. If this is the case, then it seems we have reasons to accept, that shared responsibility does not necessarily require that the individual makes a contribution to the project, she is incriminated in.

I however find it difficult to be persuaded by Bazargan's stipulation that J did not contribute, at least in some sense, and I believe it is because we implicitly accept that she does actually contribute, that we consider her liable, despite the stipulations to the contrary. Clearly, she is ineffective at her job, but we would be hard pressed to assume that the other participants (and J herself) had good reasons to think so. I.e., they were presumably more inclined to perform their roles, knowing there was a lookout. Therefore, her role is causally relevant in bringing about the heist because it contributes to sustaining the group. Further, despite what we are told, it is difficult to imagine that the mastermind would have given the go-ahead for the mission without a lookout, or that the mastermind would have been as strongly inclined. If having a lookout does not affect her decision to go forward with the heist, why would she include a lookout in the first place? The fact that she actually did include a lookout, seems to push us towards the implicit belief that having a lookout is preferable. If it is preferable, it presumably does affect her decision to an extent, i.e. it is part of the set of reasons for going forward with the plan. If reasons are causes for decision, then her role is causally relevant, i.e. it contributes to the project.

In order to test whether it indeed is our intuition that she actually does make a causal contribution, which incriminates her, we can revise the case to more clearly stipulate her role as superfluous. While doing this, we update her own, and the other participant's knowledge of her incompetence. In that case, it is doubtful that our intuitions will suggest her moral responsibility.

Imagine a mastermind hires them for a job, but she clearly states that J will not be of any help, because she will be asleep. Further, J agrees and they all have a reasonable, strong and justified belief that she will be sleeping, and J is aware that none of them are compelled in any way to go through with their job, or will

perform their jobs better, compared to a situation where she did not participate. So, if they now go through with the job, it is difficult to even accept that J should actually be considered to have the role of 'lookout', or any role in the heist. Indeed, it seems plausible that her role is in part defined by the effectiveness at performing that role, and by everyone else's acknowledgement and expectations of her effectiveness in supporting the other participants in the heist. It is not just a nametag that defines *J's role*. If we still have intuitions regarding *J's* responsibility, it's presumably rather that she fails to fulfil a general obligation to call the police knowing that the heist will commence, i.e. an omission. It is not because of any "role" she has in the ordeal aside from the contribution we implicitly accept that she makes.

Of course, Bazargan might respond that it is *J's* intention that in part defines *J's* role here. But what kind of intention would that be? It may be *J's* mere *aim* to contribute or to play a role, but since she is (let us assume) reasonably aware that she will not actually fulfil this role in any contributory sense, that seems an inconsequential notion of intention in this case. I may e.g. by writing this sentence have the aim of playing a significant role in the drought in California at the moment of writing. However, we would not claim that I play a *role* in that drought, or that I am responsible for that outcome. We would not be able to justify punishing me due to having this aim, presumably even if it would help alleviate the drought to a great extent, even saving lives. Another sense of 'intention', could be *foreseeability* i.e. I intended a certain harm, if I foresee that my action or role will foreseeably contribute to the harm. However, as stipulated in the example, going up on the roof furthers the group's goals about as much as e.g. *J* chewing some gum, tying her shoes, or something equally unrelated, i.e. not at all. This is, as stipulated, the reasonable expectation of *J*, and the expectation of the other participants. Lastly, if it is instead intention in the sense of merely *desiring* that some state of affairs obtain at some point in the future, without this implying that action is being directed at it, or is being planned to be directed at it, or that it is plausible that it is plausibly expected that it will be planned in the future, it is implausible that this would ground liability to any sanctions. Simply *wishing* for an outcome in this sense, is implausible grounds for any clear liability

or moral responsibility for an outcome *in itself*, unless we assume that wishing has some efficacy with regards to that outcome. Of course, sometimes, “wishing” can be predictive of future behaviour. If J has a deep seated desire to make a contribution to the heist, this may ground a reasonable expectation that J is the kind of person who will produce some perhaps unrelated harm at some point in the future. Perhaps such an expectation may ground J’s liability to attack now, if it will help in saving the lives of hostages.¹²⁴ However, in that theoretical case, J’s liability pertains to a different role in a different case, and not J’s role in the actual case, because, as already stipulated, J is knowingly entirely ineffective in furthering the outcome of *the heist*. Therefore, this does not support that it is J’s mere role, aside for her contribution, which grounds her responsibility.

To conclude - it is difficult to see what it is that designates J’s role and moral incrimination, if not J’s actual, potential or predictive causal contribution. If we exclude those notions of *role*, then it is not clear that there is anything left which would ground J’s moral incrimination in the case. Therefore, it does not seem to be J’s role alone which grounds her moral incrimination. Indeed, it does not seem to make sense to talk about a role, if the case lacks a contributory element.

Rounding off the observation that shared responsibility as moral membership requires that the individual performs some contributory qualifying action: All in all, it is difficult to see how an individual can share in the responsibility of a group, if she fails to make some sort of contribution to that group, and the relevant project. Therefore, it is important to incorporate some notion of contribution into an account of moral membership. What kind of contribution that should then be is still a matter of discussion which I will engage with in the latter part of this chapter.

Recapping all the conclusions

I have made a number of observations with regards to constructing an account of moral membership which I will recap in brief. 1) Moral membership is an account

¹²⁴ Though it does have certain connotations of there being some kind of “minority report”-esque thought police, which will not sit well with most people.

of shared moral responsibility where the individual is morally responsible for an outcome when that outcome may be attributed to the agent in part due to her indirect contribution to some outcome. I.e. her participation in a group project confers moral responsibility on her for the outcome of this project, beyond what is conferred on her in terms of direct moral responsibility. 2), participatory reciprocity is not necessary for moral membership. Specifically, it is not necessary for my incrimination in a collective project, that someone else has the same conception of the project, or is open towards my participation in the project. It seems sufficient in some cases that I am open towards the participation of others, and that I have a particular conception of the project we are engaged in, regardless of the others. Specifically, it requires that I alone have an appropriate degree of awareness. 3), an account of moral membership is only feasible if it can account for our intuitions of diverging moral evaluations of different participants in the group project (gradability). Lastly 4), an ascription of moral membership requires that the individual makes some contribution to the group project itself, or at least, the individual has to have a reasonable expectation that she by her action will contribute to the project in some way.

Of the above observations, the one requiring most immediate attention is the last one concerning “contribution”. In the last part of this chapter I will discuss the notion of contribution I will be incorporating into the account of moral membership.

8.3 Contribution and moral membership

In this large section, I will propose a way in which the individual can fulfil the - presumably necessary - contributory condition for moral membership. I will specify the full account of moral membership in the next section (8.4), but here briefly, roughly, specify the overall account of moral membership. This will be helpful with regards to seeing how the contributory condition fits. Firstly, assuming the individual has upstream control¹²⁵, the individual makes a certain contribution to a collective project, a project which in turn brings about the

¹²⁵ From subsection 2.4.1

relevant outcome. Secondly, the individual is aware of making this contribution, and aware of the significance of the outcome of the project. Thirdly, the outcome is morally relevant, e.g. harmful. Fourth, the cost of defection was sufficiently low. When the individual fulfils these conditions, she is a moral member of the project, and morally responsible for the outcome of that project. As should be clear, this account mirrors the direct account to a large extent. The only difference is the individual's contribution is not direct. As I will argue in this section, the specific notion of 'contribution' I will ultimately propose, which satisfies the contributory condition, is one where the individual can be said to contribute to the collective *project* that brings about the relevant outcome, when there obtains a certain sense of interdependence between the individual and the other participants. This condition is fulfilled irrespective of whether the other "participants" understand themselves as contributing to the same project. When this interdependence obtains, I will be part of what constitutes and sustains the project. When I do this, I share in the moral responsibility for the outcome of the project, again, assuming I fulfil the other conditions.

The notion of contribution I will propose builds on work by Seamus Miller (2001) and Richard Tuck (2008) respectively. Firstly, it builds on Miller's "interdependence" notion of contribution. On his account, a group of individuals can be said to contribute to a project in a morally relevant way, if the outcome depends on the group, and if the actions of the individuals within the group are interdependently intertwined in a way which sustains the project. Though this account is promising, it is not readily applicable to the voting case *as is*. Therefore, I will with the help of work by Richard Tuck, argue that the voting case can be reconceptualised in a way which does make the reasoning behind Miller's account applicable. I will first present and discuss Miller's account. Then I will propose this reconceptualization of the voting case. In the following section 8.4. I will put everything together into the full account of moral membership.

8.3.1 Interdependence as a morally significant contribution

Here I will present Miller's account. As noted, as I will show, it is not directly applicable to the voting case. However, examining it in full, will allow me to show how it can be revised and applied to the case of moral responsibility for policy outcomes.

Miller gives an account of shared responsibility which involves "collective ends, interdependence of action, and the direct and indirect contributions of the actions of the participants to the realization of those collective ends" (Miller, S. 2001: 65). Specifically, on his account, a number of individuals share in the responsibility for a harmful outcome if there is an interdependence of action in the way that "each of them intentionally performs his own individual action, but does so with the true belief that by each doing so the agents will jointly realize an end that each of them has" (ibid). This is the basic principle of his account, and the principle in light of which, I will propose how the individual can fulfil the contributory condition for moral membership.

On Miller's account, this notion of 'interdependence' occurs when two or more people perform a "contributory action on the condition, and only on the condition, that the other [contributors actually] performed theirs" (2001: 67). Miller argues that this *interdependence* can occur in at least two interesting ways with regards to shared responsibility, namely 'directly', and 'indirectly'. *Direct* interdependence occurs when all the combined acts of the participants are necessary and sufficient for bringing about the outcome. I.e., it occurs in cases where the outcome is not overdetermined, but where the threshold is exactly met for a specific outcome. This is e.g. the familiar pivotal voter type of case where – say - five people vote for a given option and four vote against, where all five were then clearly necessary and sufficient for bringing about the voting outcome, and can thus individually be considered pivotal on the condition that everyone else do their part. In this case, they all sustain the collective project through this interdependence, where if just one of them had opted out, the project would have failed.

More interestingly for this project, there is another “indirect” notion of contribution, which though it cannot be applied to the voting case in its current form, introduces an important consideration which can be incorporated into an account of moral membership. Miller’s notion of *indirect* interdependence is exemplified by cases where the outcome is actually in the relevant sequence overdetermined by the number of participants, but where the participants still participate, but only on the condition that everyone else also participates. Miller states this relationship as follows: “agent A performs action x in order to realize (collective) end e, but A does so on condition that B performs action y, C performs action z, and so on; similarly B performs y on condition A performs x, C performs z, and so on. The same goes for C and the other participating agents. Accordingly, action y causally contributes to the realization of e in part by causally contributing to the performance of x (given that the performance of x causally contributes to the realization of e). The same point holds for action x, action z, and the other actions” (2001: 72). To demonstrate this notion of indirect contribution, we can imagine a voting-case where I and four other individuals vote in favour of an outcome, but where three would have been sufficient to bring about the win. So, only two voted against. The outcome is therefore overdetermined. However, I believe, and it is true, that if none would have voted unless all five voted, then the contribution of *all* voters is necessary for the voting outcome obtaining. They all sustain the project. I would not have voted unless everyone else had voted, and the same goes for each other individual. Therefore, there is an interdependence of action. As indicated above, there is for Miller an epistemic requirement for both the direct and indirect interdependence of action. Specifically, it is a requirement that the individual actually holds a true belief that the participation of the other participants depend on her own participation, whether directly or indirectly. Interdependence is thus not just a matter of whether her beliefs about the participation of the others is reasonable. This implies that insofar as there is an interdependence of action with regards to some collective aim, then it is true that the group indeed would have fallen apart if not everyone (every single individual) in that group had contributed.

Concerning how indirect interdependence in this way could obtain, it could e.g. be when there is a formal contract between the participants of compliance. Specifically, *indirect* interdependence would e.g. obtain if everyone individually agree that they all act only on the condition that they all act (and they all have the true belief that the others would also only participate on the condition that everyone else did). It could also be less formal, where e.g. a group of people perform an action, but where it is reasonably expected and true, that if one were to “chicken out”, all would lose heart, and abandon the project. In that case, they all sustain the project by not opting out.

Miller’s notion of contribution as interdependence seems to have a lot going for it. E.g. unlike Kutz’ account, it does not incriminate individuals who clearly seem too far removed from the harmful event from being incriminated in the outcome. E.g. on Kutz’ account, the ideologically inclined cleaning lady working in the office of a general who ordered the attack on Dresden, was seemingly complicit in the outcome of Dresden’s destruction. On Miller’s account, she would be ruled out since she would obviously not have the true belief that if not all participated, then none would. Or if we recall Bazargan’s not-actual-contribution-but-still-playing-an-important-role account of contribution, the lookout J, was not an interdependent part of the heist at all. Therefore, the lookout is not incriminated on Miller’s account, which I argued was the most plausible interpretation of the scenario. If J had stayed home that day, the heist would have turned out exactly the same so there was no interdependence among the participants.

Miller gives an example to illustrate that interdependence is an incriminating notion of contribution. It is supposed to show, that interdependence incriminates the right people, and fails to incriminate everyone else. I have made many superficial amendments to the case to rule out distracting intuitions. The example is only structurally similar to Miller’s own.

Trailer killing:

Allan, Bernard and Charlie, all want to shoot and kill the innocent drunk Dean. Dean is inside his trailer¹²⁶, parked somewhere in an isolated area with no other people around than those included in the example. Dean is dead drunk, sleeping off his hangover from the day before. Dean is completely oblivious of the plans to kill him. The obvious way to kill him with a minimum amount of risk, is to shoot the gas-tank on side of the trailer, which will then explode, and destroy the trailer, with Dean in it. Allan, Bernard and Charlie all know this, and want to take advantage of it. Further, they are aware that it takes exactly two hits to make it explode (they know this, because this is a thought experiment). Bernard and Charlie do not want Allan's help for whatever reasons, and their actions are not affected by Allan's choice in the matter. They all three align their rifles, and shoot. Bernard and Charlie fulfil the conditions for indirect interdependence here, because they would not, individually, have shot if the other one had not. Therefore, since the target required two hits, it would not have obtained unless both of them shot. Allan, who "is adamant that he will make a causal contribution" (2001: 75), also shoots. They all hit at the same time, the trailer explodes killing Dean.

Now, Miller claims that only Bernard and Charlie are morally responsible for the outcome, and that Allan – *if he is responsible at all* – is only responsible in a diminished sense, compared to Bernard and Charlie. "[Allan] had the death of [Dean] as an (individual) end, and although he intentionally made a causal contribution to the realization of this end, his action was neither (directly or indirectly) causally necessary, nor was it causally sufficient" (2001: 76). Though Miller is somewhat unclear regarding what he means by diminished responsibility, it seems that he means that Allan is merely morally responsible for the action of *trying*, rather than for the *outcome* itself. The reason why Allan is not responsible for the outcome, is that Allan's action was neither causally necessary nor sufficient for the death of Dean (it took exactly two hits), because there was no indirect interdependence between all of Allan and, Bernard and Charlie. Bernard

¹²⁶ Or, RV, camper, mobile home etc.

and Charlie would have shot whether Allan had shot or not. And if either of Bernard and Charlie had given up on their project, the other one would too, which would undermine the death of Dean altogether. Each of them could thereby veto the outcome. Further, they held a true belief to that extent. Bernard and Charlie are in a manner of speaking, a compound cause of the outcome, where Allan is causally inefficacious. Therefore, on Miller's account of indirect interdependence, only Bernard and Charlie are clearly incriminated in the death of Dean, even though they from an onlooker's point of view, presumably all would seem to play an equal part.

If we share the intuition that Bernard and Charlie, but only Bernard and Charlie share in the full responsibility for the outcome, then this supports Miller's account. I.e., it supports that it is their interdependence of action which grounds their moral incrimination because that is the only clear distinguishing feature. Now, a problem with the above example is that many will be less than convinced that Allan does not have an equal share in the moral responsibility for the outcome. The reasons Miller gives for why Allan at best has diminished responsibility is that he is not an interdependent part of the group, and "simply wants to make a superfluous causal contribution" (whatever that means – how is it a causal "contribution" if it is superfluous, and he is aware of this?) (ibid). Specifically, he does not share in the collective end, because the outcome does not hinge on his participation at all (and Allan had a true belief about this). He was neither necessary nor sufficient for the outcome obtaining. Clearly a lot is going to ride on the particular beliefs the individuals has in this case, but as noted, I imagine that many will still have the intuition that Allan is indeed still partly responsible for the outcome, and for making some contribution towards it. Therefore we should examine Miller's reasons for his presumably controversial conclusion. Recall, it should be a true belief, and we can presumably add, a reasonable true belief that there obtains indirect interdependence between the participants who are morally responsible. Indeed, indirect interdependence does obtain between Bernard and Charlie. Allan is "contributing" to an outcome that is already on track to come about. Further, he is aware that he cannot produce it by himself, and he is aware that it does not require his contribution to obtain.

Further, Bernard and Charlie are aware that the outcome does not depend on Allan, because they know that nothing Allan (or anyone else) can do will affect whether Dean lives or dies. Therefore, only Bernard and Charlie are sustaining the project, and Allan is entirely superfluous. Because only Bernard and Charlie can sustain the project, only Bernard and Charlie are morally responsible for the outcome of the project.

Now, if the above reasoning is unconvincing, one reason it may be difficult to accept that Allan is not morally responsible to the same extent as Bernard and Charlie, may be that it simply is such an implausible scenario. Perhaps we have a difficult time accepting the epistemic aspect of the scenario, and thus have difficulties aligning our intuitions with it. Do we e.g. really accept that Allan, Bernard and Charlie all are certain that the gas tank only requires two hits and not three? Are they all really certain that neither Bernard nor Charlie would have shot anyway, even if one of them had abstained (e.g. what if one of them had misfired)? It is indeed quite difficult to imagine a comparable non-ideal high stakes scenario involving multiple participants, but where it would be reasonable to accept that the outcome was guaranteed in this way. And this reasonable uncertainty regarding one's causal efficacy would presumably be sufficient to incriminate the individual as a participant in the collective project. On the other hand, if we really do accept that the outcome is already given without Allan's contribution, and accept that he could really not have brought about the outcome himself, and that he has a reasonable true belief that Bernard and Charlie will hit the target or none of them will, then perhaps that will convince some that Allan is not responsible for the outcome, or at least not to the same extent as Bernard and Charlie.

8.3.2 Interdependence and voting

The question is now whether some notion of interdependence is the sort of contribution which can incriminate the individual in a group project, on an account of moral membership. So, is there an interdependence of action either directly or indirectly in the voting case? As noted earlier, we can ignore *direct*

interdependence, because the voting outcome is usually causally overdetermined. Recall, 'direct interdependence' occurs when the outcome that actually obtains, depends exactly on x number of participants, where one participant fewer would render the group act insufficient for bringing about the outcome, and where one more participant would overdetermine the outcome. So if there are exactly the necessary and sufficient number of participants, then direct interdependence occurs. Since this is – as shown in the chapters on direct moral responsibility – extremely unlikely to happen in the voting case, we should instead focus on *indirect* interdependence. Again, indirect interdependence occurs when the number of participants overdetermine the outcome, but where the participation of each, truly foreseeably depends on the participation of all. With regards to indirect interdependence, it seems obvious that in an ordinary voting case as understood thus far, indirect interdependence does not obtain either. Even in small scale voting cases, it is not true, that if not all had voted who did vote, no one would have voted. It is not even true that if not all would have voted, then only an insufficient number of voters would have voted. The reason why is obvious. There simply is no explicit agreement of compliance, or any obvious reasons to think the other voters would have acted differently if e.g. I had decided to stay at home come voting day. So Miller's notion of indirect interdependence fails to incriminate in the ordinary voting case. However, I want to argue that we can rely on a different related notion of interdependence, namely 'threshold interdependence'. When this notion of interdependence then obtains, it allows the individual to fulfil the contributory condition for moral membership. Further, the account of moral membership then allows us to hold the participants in cases similar to the voting scenario, morally responsible for the relevant outcome. Lastly, this notion of threshold interdependence also explains the intuition that Allan in the example above, has approximately an equal share of the moral responsibility.

Threshold interdependence obtains, roughly, when the agent has a reasonable expectation of being a part of the set of causes that fills the relevant threshold that brings about the relevant outcome. Further, if it is reasonable for the agent to expect to be part of this threshold, then it is also reasonable for her to

understand herself as being an interdependent part of the group and project which brings about the outcome. Again, this is the rough description, and I will elaborate the details below. This is a notion of interdependence because her action empowers others to contribute to the outcome in a significant way, and vice versa. However, for threshold interdependence to be a grounding contributory element in an account of moral membership, I need to propose a reconceptualization of the voting scenario, and cases like it. This reconceptualization was recently proposed and defended by Richard Tuck (2008), though for other purposes than holding voters morally responsible. It concerns the idea that we are allowed to understand the individual's contribution in the voting case, and cases with similar threshold features, as being a likely part of the causally efficacious set that brings about the voting outcome. This is important, because it is an alternative to how the individual could be said to understand her contribution in light of direct accounts of moral responsibility. In the next section I will briefly recapitulate the problems with accounting for the individual's direct contribution to the voting outcome as pertaining to the threshold case. After this, I will propose a reconceptualization of the voting case and tie it to Miller's notion of interdependence.

8.3.3 The problems with the threshold voting case as understood thus far:

As I argued and highlighted in the chapters on direct moral responsibility, a direct contribution in causal terms, could be understood as either being the pivotal vote, or as being part of the causally efficacious set. On the first understanding, the individual makes a direct causal contribution if she was the pivotal voter, i.e. the exact voter who filled the threshold. This only happens in the very unlikely scenario where the winning threshold was exactly met. E.g. in a simple "yes or no" referendum, if, say 50 people voted against, then I am only the pivotal voter if exactly 51 people vote in favour. The reason is, only in that scenario can the outcome be said to have directly depended on my vote. In that case, I and every other "yes-voter", taken individually, can be said to have filled the threshold, and it is true that if I had not voted, the outcome would not have come about. If

instead 52 voted in favour, then the outcome is overdetermined, because in this case, I and every other “yes-voter”, taken individually, can be said to be irrelevant to the threshold being filled. Taken individually, it is true that the outcome did not depend on my contribution, because if I had not voted, the voting outcome would still have come about. Lastly, if only 49 vote in favour of the outcome, it simply does not obtain, and thus no one who voted in favour is *causally* responsible. Therefore, the individual is almost never causally responsible. This was of course also mirrored with regards to the cognitive condition, because the individual could not be said to be aware that she would be the pivotal voter. It was simply too unlikely.

An alternative view which was presented as a possible solution to the problem of overdetermination, was that we should accept that each and every voter who is a sequential part of the threshold, is causally responsible, even if the outcome itself is overdetermined. This relied on the notion of a causally efficacious set of voters. The causally efficacious set was the precise number of votes which was actually required to fill and secure the win out of all who voted in favour of the relevant voting outcome. If e.g. 100 people vote in a simple ‘yes’ or ‘no’ referendum, and the winner won with 55 votes in favour (and 45 against), the causally efficacious set then consists of the sequence of the first 46 votes in favour, because that was the number of votes required to beat the threshold. Which 46 voters specifically may be unclear to the voters themselves and everyone else due to the anonymity of the voting process. However, as suggested earlier, it seems appropriate to say that it was the first 46 voters who cast their vote, who filled the set, even if we are unable to say which precise individuals it was. They are then causally responsible, and everyone beyond those 46 votes, are not. The problem with this was that it relied on an implausible account of causal responsibility. It relied on the idea that the particular causally efficacious set depends on the specific causal chain (compound event), which would only have obtained if they had voted. So if I had not voted “yes”, the voting outcome would still have obtained, but the compound event which included the outcome would have been different, because it would not have included “I voted yes” as a part of the causal chain. Therefore, this particular event and outcome depends on me voting “yes”. One of the main

reasons that this is an implausible account of causal responsibility, is that it is counterintuitive that we should attribute the whole outcome to the individual voter. It is simply too much of a stretch that it should be appropriate to say that the individual voter, caused the voting outcome. Further, it is also difficult to accept that the temporal sequence matters with regards to whether we can attribute the outcome to the individual. On the noted interpretation of the voting outcome, only those voters who happen to be part of the causally efficacious set, i.e. the 46 voters, are causally responsible. But again, that contradicts how we understand the voting case, or cases like it. We generally do not care if we vote first or last. Or at least, we do not care about it for reasons pertaining to causal responsibility. Again, it is mirrored in the cognitive conditions for direct moral responsibility. Though we are plausibly aware that we are part of the sequence of votes that filled the threshold, we are not aware that we are causally responsible on this interpretation. The reason is, it is not a plausible interpretation of contribution at all. It would be unreasonable to expect to bring about the whole outcome. And indeed, people by and large do not believe this. So though they are aware that they are part of the causally efficacious set, they are not aware that they have brought about the voting outcome, at least not in line with this notion of causation. Awareness pertains to belief, and since they do not believe that they actually brought about the whole outcome, they fail to fulfil the cognitive condition. Further, they ought not to believe it either.

8.3.4 Voting reconceptualization

I have now summed up the previous conceptualizations of the threshold voting scenario with regards to the direct accounts, and the problems with attributing the voting outcome to the individual voter. I will now introduce and elaborate the noted reconceptualization of the voting case, as it pertains to the moral membership account.

As already noted earlier in section 8.3, the reconceptualization can be stated like this: An individual voter contributes to the voting project, if she through her act

was likely to be part of the causally efficacious set of votes which completed that project.

Instead of the individual's contribution being understood as being the pivotal voter, or being part of the sequentially causally efficacious set of voters that fill the threshold, I will understand the voting case as a theoretical construct which closely maps the (or at least *a*) folk notion of contribution in such cases. Specifically, I will understand all votes as being cast and counted simultaneously, where we can be said to contribute to the voting project, if it was likely that our vote was part of the causally efficacious set of votes that brought about the voting outcome. I refer to this as "threshold interdependence". The reason is - as I will argue - there is an interdependence of causes implied when we understand our contribution as being a likely part of the causally efficacious set in this way. Further, and as I will argue later on, this interdependence is what makes it appropriate for us to say of the individual that she contributed to the particular project. Threshold interdependence incorporates an observation highlighted in previous chapters. The observation is that we do not seem to care whether our votes are cast first or last. We do not seem to care about the voting sequence at all in the ordinary case. This suggests that we consider all votes cast in favour of the winner to be equally relevant. This reconceptualization is not an attempt to provide an alternative account of how the individual can make an actual causal contribution to the voting outcome itself. Rather, it is an account of how the individual plausibly can be said to make a normatively relevant contribution to the group project, in a way which makes her qualify for membership of the particular project. The project, which is then the *direct* cause of the outcome.

Allow me to restate and elaborate the reconceptualization: An individual voter contributes to the voting outcome if she through her act was likely to be part of the causally efficacious set of votes that brought about the outcome that actually comes about. Specifically, if we assume there is a random distribution of votes for the relevant candidate, then we can ascribe a particular likelihood that a specific vote would be included in the causally efficacious set. We can then from the size of the pool of voters who voted for the relevant candidate, and the size of the

causally efficacious set vs. the size of the set of inefficacious voters determine the likelihood that any given voter was part of the causally efficacious set. The criteria for determining whether a voter is a part of the causally efficacious set, is simply random distribution. As seen in cases where the participants usually do not know who will be part of the causally efficacious set (e.g. voting and firing squad cases), the actual causal succession of the participants is appropriately understood as random selection in the above sense. When it was likely that I was part of the causally efficacious set, my action contributes to the project, because it empowers the other contributors to bring about the outcome (and vice versa). It empowers them, even if it does not contribute to the *outcome* directly in any strict causal sense. The way the individual then contributes, is by performing an action which results in her being a threshold interdependent part of the voting project. When she does this, she helps constitute and sustain the project, 'constitute', and 'sustain' in a sense at least.¹²⁷ She does not necessarily make a direct contribution to the policy *outcome* through our voting action, but by being a likely part of the set of causes that brings the *project* to fruition, she is contributing to that *project*. Recall the firing squad cases from the previous chapter. The bullet of the shooter who is behind the one way mirror, may or may not be a relevant cause in the explosion of the barrel which kills the innocent prisoner. The example did not specify it. But, because she knew it was a *likely part* of the set of causes, she is helping sustain the project which causes the death of the prisoner. Therefore, we would consider her morally responsible even when her bullet happens to arrive last, and is thus causally irrelevant (subsection 7.4.2). In a sense, it could have been her bullet, and it could have been her vote that was

¹²⁷ I am using 'constitute' and 'sustain' in a slightly technical sense. The intuitively clearest examples of someone constituting and sustaining a project, will presumably be ones where the participating individual is actually clearly causally efficacious. But, what I have shown with various examples, most clearly the firing squad cases, is that our intuitions support the notion that we actually are contributing in some cases, even if we are not actually causally efficacious in bringing about the relevant outcome. The thing that grounds our contribution in these cases seems instead to be the fact that it could just as easily have been us who were part of the causally efficacious set, as those who actually were part of it. Indeed, my participation, even if I am not part of the relevant set, increases the chance that the project will be successful. Therefore – assuming I am aware of my part in the project, or should have been aware - I am contributing in the sense of helping constitute and sustain the project. I was just as likely to be an actual building block in its succession as anyone else, and it would have been closer to failure if I had opted out, instead of participating.

part of the causally efficacious set. All this is of course under the assumption that the shooters at the moment of firing their rifle, lack insight with regards to the actual sequence. If I e.g. knew that my bullet would arrive last, then this awareness presumably changes the verdict of the case. In that case, I would be aware that I was not sustaining the project. But only in that case. The problem is, that the *likelihood*, i.e. the expected probability distribution, will be screened off by our knowledge of the actual facts of the matter. It is not likely that we will be part of the causally efficacious set, if we know we are not. The notion of threshold interdependence is thus a normative notion, and it is predicated on our epistemic circumstances about the likelihood of being part of the causally efficacious set, rather than the *actual* external causal circumstances of the scenario.

If threshold interdependence is a normatively relevant notion of contribution which can incriminate the individual in a collective project, the next question is then whether the ordinary voter is actually a threshold interdependent part of the project. Indeed, in the ordinary voting case, it is quite likely that we do make a contribution to the group project in ordinary voting cases, at least insofar as our preferred voting outcome wins out. It is indeed quite likely that our vote was part of the causally efficacious set in case the preferred outcome obtains. In the normal voting case, due to the competitive vote seeking nature of political candidates, the race is usually going to be very close. Therefore, the probability that one's vote was part of the causally efficacious set, is overwhelmingly high. If e.g. 180 million people vote, and the winner receives 100 million votes, and the loser then receives 80 million votes, then the causally efficacious set is the amount of votes it took to beat the threshold, namely $80 \text{ million} + 1$ votes. Assuming a random distribution, the likelihood that my particular vote was part of the $80 \text{ million} + 1$, and not the $20 \text{ million} - 1$ wasted votes is approximately 75% i.e. quite likely compared to the chance of being the pivotal vote. It is therefore a reasonable expectation that the voter is part of this set. Obviously, the likelihood of being part of the causally efficacious set does not change when we multiply the total number of votes, as long as the ratio is kept fixed. Though I have not been specific regarding what should count as "likely" here, it is not a significant stretch to claim that anything above 50% is indeed likely.

Regarding ‘threshold interdependence’, and why it can be understood as interdependence, recall Seamus Miller’s broadest description of the notion. On his account, what interdependence amounts to is: “each performed his contributory action on the condition, and only on the condition, that the other(s) performed theirs” (Miller, S. 2011: 67). Presumably, many voters only vote if they believe a sufficient number of others also vote. But as I have argued, we need to reinterpret interdependence. What is important is not whether I vote because others vote. It is whether I vote while foreseeing that my vote will sustain the project. Again, this is not an account of direct individual causation, or causal responsibility. Rather, it is an account of the conditions under which voters understand, and ought to understand, themselves participating in a particular project. If they participate under these conditions, they can be said to contribute to the project. They participate whether or not their vote actually made a difference to the outcome obtaining. When they participate, they – to use a phrase from an earlier chapter – test their mettle. They act on their willingness to sustain the project. If the project foreseeably brings about a morally relevant outcome, this tells us something about who they are. It allows us to reevaluate their moral ledger. There may be other ways to contribute, and other reasons some people vote, but the claim is here simply that this particular notion of contribution does constitute a qualifying action for moral incrimination in the relevant project i.e. moral membership of that project. Overdetermination, i.e. the fact that more than enough will usually vote for my preferred candidate, is not a problem here. The issue is not whether I actually make a direct causal contribution to the outcome. The issue is whether my participation expectedly depends on the participation of others, and vice versa for them, whether their participation depends on my participation. If it is likely (which it is) that I am part of the causally efficacious set, and if their participation depends on the causally efficacious set being filled, then their participation is likely to depend on my participation (and vice versa¹²⁸) as well.

¹²⁸ There is for all individuals, to use the previous example, a 75% chance that they will be part of the causally efficacious set assuming a random distribution.

Now, this notion of interdependence is quite different from the notion Miller proposes. On his account, indirect interdependence meant that the agents in question could each veto the outcome by defecting, because there was e.g. an agreement that none would participate unless all participated. Obviously however, this notion of interdependence applies to a limited range of cases. Therefore, his notion of interdependence is simply too demanding for many collective actions cases. It would e.g. be too demanding for the firing squad case.¹²⁹ I propose instead that the notion of interdependence I have here suggested is sufficient for moral incrimination in a collective project, assuming the individual fulfils certain other conditions. If applied to the shooter case of Allan, Bernard and Charlie, it is then plausibly the case that Bernard and Charlie, even if they would not have gone ahead with the project if either one of the two had abandoned the project, and even if they were hell bent on ignoring Allan's contribution to the outcome, they cannot escape the awareness that they are all part of the causally efficacious set in the actual sequence where they actually act. They have a reasonable expectation that they through their actions will be a likely part of the causally efficacious set that brings about the explosion that kills the victim. Given the setup of the scenario, there is a 2/3 chance that each of them will be part of that set, which certainly counts as 'likely'. They are aware of this, and they still chose to act. Therefore, they are sustaining the project of killing Dean, whether they want to include Allan or not. Their moral incrimination is of course not surprising and they were already morally responsible on Miller's account. More importantly, Allan is also incriminated. When he actually acts, even knowing that Bernard and Charlie want to be acting alone, and even though he simply wants to make a superfluous causal contribution, he cannot escape that he is a likely part of the set of causes that killed Dean, or that he is actually sustaining the project. Thus they all ought to recognize that an interdependent relationship obtains between them, i.e. that the outcome in all likelihood depends on the contribution of each of them.

¹²⁹ I would e.g. shoot even if the number of shooters clearly overdetermined the outcome, and even if not all would have shot.

But what if a fourth, a fifth, or a sixth etc. person were also shooting? Then with the explosion-threshold being two hits, we would then have to acknowledge that they are each less likely to be part of the causally efficacious set. So if there are six shooters, it will be $\approx 33\%$ chance that a given individual shooter, was part of the causally efficacious set. What would the implications be when we continue to lower the probability that the participants are part of the causally efficacious set? I can think of at least two possible verdicts which possibly has implications with regards to an observation from earlier, which concern *gradability*.

1) If the likelihood of being part of the set becomes too small, the participants are **excused for their participation** – i.e. it was simply too insignificant a contribution (assuming they were also aware of this). Presumably it would take a lot more shooters than six for us to consider their contribution to the project *too* insignificant for a negative moral evaluation. The likelihood of them being part of the causally efficacious set is still much too high in light of the severity of the outcome of the project, and we would plausibly still hold that they were sustaining project. But perhaps at some point, there is a point where they each become too unlikely to participate, that we should accept that they are simply excused because they are not plausibly sustaining the project anymore. Then the chance of belonging to that threshold is then too unlikely, and we would consider the contribution to the group project negligible.

2), **they are less accountable for the outcome** to a degree relative to the likelihood that they are part of the causally efficacious set, i.e. less accountable to the degree that the outcome depended on their contribution (again, assuming they are also to some extent aware of the likelihood). So in a case where it is expected to be extremely likely that they will be part of the set that actually brings about the outcome, they will be deserving of significantly harsher condemnation. The important thing to note is, that if we assume 2), we can actually account for our intuitions of gradability, because that would then be a function of the strength of our contribution to the group project, where the smaller the likelihood of being part of that set, would constitute a smaller contribution to it. It is difficult to gauge the relationship between likelihood of

being part of the causally efficacious set and the accountability of the individual. As we saw in the shooter case from the previous chapter, when the outcome of the project is such a significant harm as killing an innocent person, we would presumably have to add an enormous amount of shooters, before our intuitions of blame and punishment start decreasing along with the decrease in the likelihood that we are part of the causally efficacious set.

Summing up this substantial section: I have introduced an account of how the individual can be said to fulfil the necessary contributory condition for moral membership. In the voting case specifically, an individual voter contributes to the voting project, if she through her act was likely to be part of the causally efficacious set of votes that brought about the success of the project. In order to arrive at this account of contribution, I firstly introduced David Miller's interdependence account, where interdependence obtained among the participating individuals, if the individuals performed their contributory action on the condition that others perform theirs. I then examined it, and argued that though interdependence seemed to be a morally significant notion of contribution, Miller's notion of it was too demanding for the voting case. On his account, the individual is morally incriminated in a group project, if that individual was either a directly interdependent part of the project, or an indirect interdependent part. *Direct* interdependence occurs in the case where the individual is *the* pivotal voter. *Indirect* interdependence occurs when the project hinged on the participation of all participants, where if not all had participated, none would. Since these notions of interdependence were too demanding, I introduced the notion of *threshold interdependence* which relied on a reconceptualization of the voting case. Instead of understanding the voting case in terms of the individual making a direct contribution to the voting outcome as I had done earlier, I argued that an alternative way of perceiving it was as a project. Here the participants contribute to voting project, if they were likely to be part of the causally efficacious set that brought about the success of the project. They are sustaining the project, even if they are not making a direct causal contribution to the voting project itself, as long as it was likely that they were part of the causally

efficacious set. In the next section, I will put this notion of contribution into the full account of moral membership.

8.4 Moral membership, the full account

It is now time to spell out the full account of moral membership which is sufficient for individual moral responsibility in cases like the voting case. I will present the individually necessary and jointly sufficient conditions, then clarify them.

I am a moral member of a collective project that brings about an outcome O, iff:

- 1) I help constitute or sustain the project that brings about O, in the right way.¹³⁰**
 - a. E.g. there obtains threshold interdependence between myself and the other participants.
- 2) I fulfil a cognitive condition:**
 - a. I am aware that I am constituting or sustaining the project that brings about O (or I ought to be aware)
 - b. I am aware of the significance of O (or I ought to be aware)
 - c. I am sensitive to moral considerations
- 3) The outcome O is morally relevant.**
- 4) Defection was relatively costless.**
- 5) I have upstream control:**
 - a. I fulfil some freedom relevant condition
 - b. I am not physically forced
 - c. I am not under strong abnormal psychological compulsion
 - d. I am not strongly manipulated

I will now elaborate the conditions.

¹³⁰ I do not want to rule out that there are a multitude of ways in which the individual can contribute to a given project. I am simply giving a positive account of one way in which the individual can make a morally incriminating contribution. The example specified under 1)a is the particular account I have presented in this chapter. On the other hand, I do not want to make the claim that any sort of constituting or sustaining a project is necessarily incriminating.

Condition 1) - I help constitute or sustain the project that brings about O, in the right way, is unspecific in its formulation. The reason is that I do not want to rule out other ways in which the individual can be said to contribute to the relevant project in a morally incriminating fashion. However, as stated in 1)a, the claim is that the individual can indeed fulfil this contributory condition in the specified manner. In order for there to be threshold interdependence, outcome O has to come about as a consequence of a specific set of causes filling a threshold. Further, I have to perform an action which will likely result in me being part of the specific set of causes that brings about the project, that brings about the outcome O. Regarding set individuation, I may be a member of a near infinite number of sets, some of which can be said to cause the project, and the outcome in question. E.g. I am a biped, and if the voters in question are all bipeds, then I am a member of the set that brings about the outcome just by being a biped. But obviously this is not the *relevant* set. The relevant set is the specific set of voters who along with myself, voted in favour of the winning candidate.

Condition 2) - I fulfil a cognitive condition. It mirrors the cognitive condition discussed in chapter 4. An outcome is only attributable to me, if I am aware of the relevant facts, or ought to be aware of them. If I vote, but fail to understand that my action will indeed make it likely that I sustain the collective project then – unless I am culpably ignorant – the outcome of the project will not be attributable to me. Further, the same applies if I am unaware, or have a false belief regarding what the significance of the outcome is. If I believe that the ultimate outcome of the collective project is $\neg O$, but it really is O, then O is not attributable to me, unless again, I ought to know. Note, I can fulfil this condition even if the other participants are completely ignorant of my participation, and even fail to include my participation in their estimation of the probability that they will be part of the causally efficacious set. Specifically, participatory reciprocity is not a requirement of my moral incrimination at all. Indeed, the other participants can be unaware of what the project entails, and indeed have a wholly different conception of the project than me. This does not affect *my* moral incrimination. Nothing in any of the conditions rule out that the other participants can be entirely opposed to my contribution in the project.

Lastly, this fulfilment of this condition also requires that I am sensitive to moral considerations. This relates to the agent's ability to recognize something as a moral reason, something that may e.g. be undermined by mental illness or other mental abnormalities. The stereotypical psychopath would be an example of someone failing to fulfil this condition. She may be aware that some action can produce some outcome, but she is not necessarily aware that it is morally significant. She e.g. does not recognize the potential harm that comes about, as a relevant moral reason. It is a contentious issue whether moral competence undermines moral responsibility (e.g. Levy 2013)¹³¹, so the inclusion of this as a necessary condition will of course have to depend on whether it indeed plausibly undermines accountability. However, I will not be engaging with that particular discussion, but simply include the condition, while noting the contingency of it.

Condition 3) - The outcome is morally relevant, specifies that the outcome has to be morally relevant. This in part mirrors the attributability/accountability distinction from the direct account of moral responsibility. We can attribute a morally neutral outcome to a person, without this making the person accountable, i.e. an apt target for our reactive attitudes or punishment. Only if the outcome is morally relevant can they be accountable for it, e.g. deserving of blame and punishment. Moral membership is an account of full-fledged indirect moral responsibility. Moral responsibility requires both attributability and accountability. Therefore, a necessary condition is that the outcome is morally relevant. The outcome can e.g. be morally relevant if it results in a significant harm that would not have obtained if the outcome had not come about.

Condition 4) - Defection was relatively costless. Specifically, if e.g. I am being coerced into contributing, this will usually at the very least mitigate my accountability. Or, if my livelihood, or my loved ones' lives depended on my contribution, it would presumably also mitigate or exculpate. Of course, the way the condition is stated, it seems to imply that my moral membership can be

¹³¹ As noted earlier, it is a contentious issue which I will not engage with, whether lack of moral competence in this sense actually would undermine moral responsibility.

undermined if there is a large financial loss involved in my participation. E.g. if I stand to gain a lot of money by participating, there can be a large relative cost incurred if I abstain. However, that this would undermine my moral membership seems implausible to most. In cases of direct moral responsibility, the hitman who receives a lot of money for shooting someone, is not excused simply because he in relative terms, would have lost money by *not* taking the assignment. Therefore, it should be clear that the whether the particular “cost” is relevant, will depend on what that cost is.

Condition 5) - I have upstream control. I have already specified that I will not try to defend any particular account of how the individual fulfils this condition. I will simply assume that the individual is not constrained in a moral responsibility undermining way due to e.g. determinism, or because she is physically forced to contribute to the relevant project.

If the individual fulfils these conditions, then she is a moral member of the project, and morally responsible for the outcome that follows from that project. An important thing to note is that nothing in the above conditions require that the other participants are also morally responsible. Having a shared intention (multiple agents in the set having the same intention to bring about the outcome, or others participating in a cooperative project) is not necessary for *my* moral membership in the project. The other participants may be wholly ignorant of the fact that they are contributing to the outcome, without this affecting whether I am morally responsible. Moral membership is individualistic to the extent that it does not require anything but the structural organization and causal force of the other participants. Further, moral membership does not require that the group in question is a full-fledged corporate agent. There does e.g. not have to be a group organisation that lasts over time, or an internal decision-making structure. It does not matter why the *other* agents act, and there does not need to be any epistemic interdependence in the sense Miller proposed where they have a true belief that defection will veto the project. Rather, what is important is that *my* conception of the project and the outcome of it, corresponds to the actual facts of the matter, and that *I* indeed am a likely part of the causally efficacious set. My moral

incrimination does not depend on whether a given case can be construed as a joint action. All other participants can be acting for different motives, or think they are participating in a different project altogether. They can be entirely oblivious to my participation, and could even actively try to undermine my participation in it, without this undermining my moral membership. Roughly, I am a moral member of a given project, and thus morally responsible for the outcome of that project, if I help constitute or sustain that project in the right way, by being a likely part of the causally efficacious set that makes the project succeed, while being aware of my role in this.

Something that may constitute a limitation to this moral membership account as presented so far, is that it is only applicable to cases which can be construed as threshold cases. But some intuitive cases of shared responsibility are not threshold cases in this sense. This is of course not a problem in itself, because this is exactly the type of case we are interested in, namely the voting case. In non-threshold cases of shared responsibility, such as e.g. the kids throwing rocks (subsection 7.4.2) in the previous chapter, it is difficult to construe that case as a matter of a threshold being filled for producing a given outcome. It is only one of the participants who actually happen to bring about the outcome, compared to the threshold case, where we are allowed to reasonably perceive of our contribution as an equal part of the thing that brings about the outcome. In those other types of cases, perhaps it is more appropriate to construe the notion of interdependence differently. In that case, it is presumably more appropriately stated as the perceived increase in the likelihood that the group will produce the outcome, that is brought about, which makes me a moral member of that group. So if I am aware that the likelihood of the other kids hitting the window (though they are actually trying to hit the Frisbee), is 50%. But I am also aware that if I join in, the likelihood of that outcome obtaining is going to be 60% (let us assume). In this case, I am helping sustain the project in a different sense.¹³² I will however not attempt to develop or defend this here.

¹³² Peter Vallentyne (2008) has argued for an account of *direct* moral responsibility where we can be morally responsible for increases in probabilities that outcomes obtain which can possibly also be

8.4.1 Testing and critiquing this moral membership

In order to test this account, I will examine Kutz main test case from the previous chapter, i.e. the Dresden case, to see whether it generates plausible verdicts in this case. The Dresden case, as Kutz presents it, is not a clear threshold case. So it needs to be reconstrued. It needs to be a case where the whole destruction is contingent on a certain number of bombs filling the threshold, where below this, no destruction obtains. As noted in the previous chapter, the way Kutz proposed the case, it seems that we could potentially explain the moral incrimination of each individual airman, by reference to their direct contribution to the outcome. So as stated, it was not necessarily a case that clearly required an account of shared responsibility to justify the moral incrimination of the individual participants. Therefore, the following example is a better test case than his original one, because the individual does not necessarily make a direct contribution, due to the threshold structure of the setup.

Dresden force field:

Dresden is protected by a force field, and it takes, let's say, exactly 8,000 bombs to overload it, where no harm is caused by fewer hits than that, and where more bombs fail to add to the outcome (e.g. these are a special kind of anti-force field bomb which only damages force fields). If it overloads however, it will cause an electrical surge which has the foreseeable effect of a massive city wide lightning strike, which will produce an enormous amount of harm to Dresden and kill many of the inhabitants. Imagine then that there are, say, 10,000 bombers, each having one bomb. Switching to the second personal perspective, imagine that you are aware that statistically one in 10 bombs will fail to detonate. You are also aware that there are approximately 10,000 other bombers aside from you participating in the booming run, and that they all will hit (it is a big and easy target, and there are no flak cannons, or any other

applied to the collective context. Mellema, G. has also argued for the notion that we can be morally responsible for risks, understood as likelihoods that harmful outcomes obtain (1987).

defensive measures to disrupt the bombing run). Lastly, in this video game-esque scenario, you are indeed aware that it takes exactly 8,000 hits to destroy the force field (and this is true). With all this in mind, you fly your plane over Dresden and release your bomb. You do this while having good reasons to expect that it is likely that it will hit and detonate, and produce the effect in conjunction with the other participants. Further, no one is *forcing* you to do this, and there is no pressure on you to go through with the task aside from your own reasons for causing this harm. Lastly, you were dishonourably discharged from the army the week before, and you are not allowed to participate, and none of the other participants are aware of your participation. You just snuck into the air base, and stole a plane, because you really wanted to participate. Indeed, everyone else is under strict orders to hinder your participation, and ultimately shoot you down, if they see you violating your discharge (and you are aware of this). But again, they do not recognize you, and you are aware that you will not be spotted. In the end, you hit the force field, and it is overloaded, and many innocent people die.

Now, are you morally responsible for the outcome, and can the moral membership account explain this? I am going to imagine that the case is clear enough to give us the intuition that yes, you are incriminated in the harm at least to the same extent as every other bomber. The outcome (or part of it) is attributable to you, and you are accountable for it. I am going to assume we have a clear intuition regarding this, even if we do not know whether your bomb actually detonated. Further, it seems clear that any *after the fact* knowledge of this is not going to change our intuitions. This is good, because this indicates that our intuitions are not tracking a direct notion of moral responsibility, which presupposes direct causal responsibility. Further, assuming you have upstream control and you are sensitive to moral considerations, you fulfil the conditions for moral membership. Threshold interdependence obtains, and you are *aware* of this. The outcome *is* morally relevant, and the cost of defection is extremely low (in fact, it costly to participate, since the pilot risks getting shot down). So it seems everything checks out according to the moral membership account. On the other

hand, e.g. Kutz' account of joint action would have difficulties explaining this, e.g. due to the adversarial attitude of the other pilots, given they would actually obstruct your participation if they could. Further, Miller's account does not apply either, because there is neither direct, nor indirect interdependence in the sense he imagined. None of us can veto the outcome. Further, my contribution has no direct effect on the outcome obtaining, so I am not morally responsible in light of a direct account of moral responsibility.

Now, it seems the moral membership account does incriminate those individuals which are intuitively accountable for the outcome in question. However, is it sufficiently exclusive, not to include anyone who is intuitively not morally responsible for the outcome? Does it e.g. incriminate the ideologically inclined cleaning lady? (section 7.5) Is she a moral member? A plausible account should not include her on par with the other participants, at least if it is going to be in line with our intuitions of moral incrimination. As the case was stipulated, she took the job in order to further the war effort. Further, she actually wanted the destruction of Dresden, and that's why she was cleaning his office (as a way of helping the war effort). The question then is whether she is a moral member. So, does she perform an action which she foresees is likely to make her part of the causally efficacious set that brings about the outcome? It does not seem to be the case. The causally efficacious set that brings about the outcome pertains to the number of bombs necessary for bringing about this destruction. It would be unreasonable to expect that she by her action is likely to be part of that set. In this case it seems appropriate to apply direct accounts of moral responsibility, and examine whether the act of the general depends on her contribution, and whether she was aware that she had this impact on him. But she does not have any reason to think so, so she should not be considered directly morally responsible either. Even though she seemed to fulfil the conditions for joint action in on Kutz' account, and was implausibly incriminated on his account, she is not a moral member, which is in line with our intuitions. Indeed, threshold interdependence is a quite demanding, even though it pertains to a limited range of cases, i.e. threshold cases. So it is unlikely that someone will be incriminated who is not clearly morally responsible.

8.5 Summing up this chapter

In this chapter, I have proposed an account of shared responsibility which I refer to as moral membership. When the individual fulfils the conditions for moral membership, she is morally responsible for the outcome of the project of which she is a moral member, or part of that outcome. The outcome is attributable to her, and holding her accountable is thus justified. The degree to which she is accountable, is presumably related to the likelihood that she is part of the causally efficacious set, though I have not attempted to defend this claim in full.

This account builds on the investigation from the previous chapters on direct and indirect moral responsibility. I started by highlighting what I considered four important observations from the previous chapters, pertinent to the construction of an account which could give a positive verdict in the voting case. These observations were 1), a successful account of moral membership should not require that the individual makes a significant direct contribution to the relevant outcome itself. 2) I can be a moral member of a project, even if there is no 'participatory reciprocity' between myself and the other participants. 3) an account of moral membership ought to be able to account for our intuitions of varying degrees of accountability in relevant cases. And, 4) shared responsibility as moral membership requires that the individual performs some contributory qualifying action. Based on these observations, I then constructed the account of moral membership. I proposed that the individual becomes a moral member when a certain interdependence obtains between herself and the other participants, where she helps constitute or sustain the project. I specified that she can be said to sustain such a project, when it is likely that she is part of the causally efficacious set, that makes the project succeed, even if she fails to make a direct causal contribution to the outcome of the project itself. I then applied the account to Kutz' "Dresden" case, and illustrated that the account is plausibly at the heart of our intuition of moral incrimination in that case.

I will in the next chapter, apply this account to the voting case specifically. I will show that being a moral member of the project that produces the relevant policy

outcome, does incriminate the individual. Indeed, I will show that this incrimination can ground quite severe accountability for that outcome.

9 Conclusion

9.1 Introduction

In this chapter, I will firstly in brief sum up the previous chapters, and elaborate how the overall question set out in chapter 1, has been answered. After this, I will apply the moral membership account of indirect moral responsibility to the case of a just war. In doing this, I will be able to show how the moral membership account can affect the conclusions we should draw with regards to citizens' accountability, and thereby their liability in war. As noted in chapter 1, an important criterion for an account of moral responsibility is that it should make citizens take their political responsibility more seriously. If a link can be demonstrated between the individual citizen's voting action, and her liability in war, this should lead people to accept the severity of the responsibility they are taking upon themselves when they cast their vote. This should in turn ideally affect willingness to acquire information, and deliberate on the relevant issues, and make a more informed choice in the matter. Specifically, I will show that their moral membership will affect their liability to attack in war. If it is clear, that it may ultimately be morally permissible to kill someone for voting for a given candidate, this should certainly lead to some reflection about the seriousness of the voting act.¹³³ After having applied it to the voting case, I will conclude on the whole project.

9.2 Summing up the previous analysis

I have argued that individual citizens can be morally responsible for policy outcomes, if they are moral members of the collective project that brings about those outcomes. Roughly, they are moral members of a project, and the relevant outcomes, if they were aware that their contribution was a likely part of the

¹³³ Of course, unless it would simply be considered a *reductio ad absurdum*.

causally efficacious set that actually brought about the relevant project that caused the policy outcome. In order to arrive at this account, I firstly examined whether the individual voter was morally responsible in light of making a direct contribution to the voting outcome. I argued that this was implausible. The main problem was that it was not possible to give an account of how we should attribute the policy outcome to the individual, in light of her direct contribution to that outcome. Particularly, she was not clearly causally responsible. The problem was that the outcome did not depend on her direct contribution in any clear way. When understood as a threshold case, the contribution of the individual is almost always insufficient or superfluous. Further, this was mirrored with regards to the cognitive condition also. Specifically, she was not aware that she would make a direct contribution even if she did, and she was excused for her ignorance. If either condition – both causal responsibility and the cognitive condition – is necessary for an ascription of direct moral responsibility, such an account fails in the voting case. We do not make a direct contribution, and the outcome cannot be directly attributed to us.

In light of the failure of the direct account, I then examined indirect accounts of moral responsibility. Here the most promising one was Kutz' complicity account. On Kutz' account, an individual is complicit in a harm, if she engaged in a joint action that brought about that harm. Unfortunately, the voting act is not clearly a case of joint action, even if joint action is sufficient for moral responsibility. The problem is the disorganized actions e.g. in light of the various diverging motives of the voters, which make it difficult to construe their activity as a joint action proper. Two individuals may vote for the same candidate, but do so having contradictory ends, or contradictory conceptions of what the outcome of the candidate winning will be. For this reason, I argued for loosening the conditions for joint action. Doing this however, means that the activity the individual participants are engaged in, presumably can no longer be characterized as a joint action proper. Instead of joint action being the basis for attributing the relevant outcome to the individual, I argued that an individual action, which is performed which is likely to be an interdependent part of the set of causes that brings about a project, is a sufficient notion of contribution to incriminate the individual in this

project. When that project then is the cause of a harmful outcome, that outcome can then be attributed to the individual. It can be attributed to the individual regardless of whether the other individuals shared her conception of the project, and regardless of their openness towards her participation in her conception of the project.

The only thing missing, is the actual application of the moral membership account to the case at hand. So in this concluding chapter, I will apply it to the voting case. In doing this, I will apply the discussion of moral responsibility for policy outcomes, to the discussion of non-combatant liability in war.

9.3 Moral membership and non-combatant liability in just war

In this section, I apply the moral membership account to the context of non-combatant liability in a just war (actually an unjust war) scenario. I will show that moral membership can ground moral responsibility for policy outcomes. If an argument can be made, that runs from the moral membership of the individual, to her liability to getting killed in war, and if this conclusion is intuitively plausible, this will then support the plausibility that the moral membership account indeed grounds moral responsibility for policy outcomes. It will support the moral membership account, because the conclusion that civilians are liable to attack, goes against our pre-theoretic judgments. The moral membership account has to be able to lift a significant burden if it is going to convince us otherwise. I will assume that most people share a strong intuition that civilians are never valid targets of attack in war. There may be different reasons for this, but I will assume that part of this intuition is due to us not considering them clearly morally responsible *for* the war. This allows me to test this intuition by seeing whether an account of moral responsibility can affect our verdict regarding the citizen's liability. The strategy is to construct an argument for citizen liability to defensive killing, which includes their moral membership as a central premise. If the moral membership account can lift the burden and lead us to the intuition that ordinary citizens can be liable to attack compared to a scenario where she is not clearly a moral member, this will provide support for the plausibility of the account.

Further, it will also provide an example of one kind of sanction that may be appropriate in light of individual moral responsibility for policy outcomes.

The intuition that civilians are not valid targets in war has some foundation both in and outside academia. On the outside, it is usually assumed that people by and large consider attacks on civilians impermissible. This is also supported by data (Gallup.com 2010). Further, it is also codified in the Geneva Convention (IV) regarding the Protection of Civilian Persons in Time of War (1949), which prohibits attacks directed at the civilian population. Further, in recent years, many hold that we should have extreme restraint when acting in ways which has foreseeable civilian casualties (e.g. Walzer, M. 1977, ¹³⁴ Primoratz, I. 2010), to the extent that direct attacks on civilians usually cannot be practically justified. However, if we accept that war can be a form of policy, and if war represents the will of the people in the relevant sense, and if this will is represented in the policy of going to war, and if this will, as I have argued, is constitutive of their moral responsibility, then it seems, tentatively, that some civilians could be valid targets in war. This latter way of understanding civilian liability in war, is referred to as the *reductivist* view¹³⁵ of non-combat liability. The reductivist view assumes that the judgments we ordinarily apply in the context of ordinary society, should be the basis for our judgments in war. A non-reductivist view holds the opposite, that war holds special normative significance, and requires a special set of rules with regards to e.g. noncombatant immunity. The non-reductivist is usually attributed most famously to Michael Walzer, who claims that the two central concepts of *jus ad bellum* and *jus in bello* i.e. ‘right to war/fight’ and ‘right in war’ are – in Walzer’s words – “*logically independent*” (1977: 21). This means, that the rules governing morally legitimate actions in war, do not necessarily depend on the particular reasons or justification for engaging in war, and vice versa. “It is perfectly possible for a just war to be fought unjustly and for an unjust war to be fought in strict accordance with the rules”(ibid). Reductivists, most famously Jeff McMahan, hold

¹³⁴ I should note that it has been argued that Walzer’s view makes provisions for an exception of this principle in the case of the supreme emergency (Nathanson, S. 2010: 139).

¹³⁵ E.g. Jonathan Parry uses this label (2014).

a different view, where the rules for conduct in war ought to be an extension of the rules we usually associate with society at large. So when e.g. a “murderer is in the process of killing a number of innocent people and the only way to stop him is to kill him, the police officer who takes aim to shoot him does not thereby make herself morally liable to defensive action” (McMahan 2009: 14). The reason the officer does not become liable herself is that one party is an *unjust* aggressor, who through her aggression has made herself liable to attack. She has thereby lost her right not to be harmed. One party is justified because the other party constitutes an unjust threat. Reductivists claim that the considerations which pertain to such a situation, can also pertain to the war scenario. If one party e.g. invades another party unjustly, they may thereby have made themselves liable to attack. The other party, the party that is being invaded, is not necessarily also liable to attack. Further, and this is the important part, if the citizen is morally responsible for the policy outcome of the unjust war, it may be possible to establish an in principle case for her liability to defensive killing. This would not necessarily have been permissible on Walzer’s non-reductivist account, because the rules, e.g. the rules governing the citizens’ liability, do not necessarily depend on concerns pertaining to whether they are morally responsible for the war. In Walzer’s case, citizens are simply not valid targets under the internal logic of war (1977: 21).

In this chapter, I will apply the developments from my project to the reductivist view. This means that I will be assuming that the moral status of non-combatants, and whether they are liable to attack in war, will depend on whether they have done something to lose their right not to be attacked. The test of the moral membership account is then whether a non-combatant citizen who qualifies for moral membership, can make herself liable to attack in war. I am not taking any stand on which interpretation is more plausible, the reductivist or non-reductivist. I will simply be *assuming* without further argument, that McMahan is right when he states: the “*liability to attack in war is moral responsibility for a wrong that may be permissibly prevented or corrected by means of war, or by an act of war*” (McMahan, J. 2009: 204). Here that “wrong” is being a moral member of a project which produces a certain harm. This is important to this project, because “the claim that civilians are not legitimate targets in war should follow from the claim

that they are not responsible for a wrongful war or for the wrongful acts of which it is composed” (ibid). If my account of moral membership is correct, then insofar as the individual participant is a moral member, there may be situations where she is liable to attack, even though she is an ordinary citizen as described in this project. A note before producing the argument: we should as suggested¹³⁶ be sure to distinguish between the morality of war, and the legal justification for acts in war. It may seem morally permissible, to target civilians in war. But it may still be true, that deliberately attacking civilians should be legally indefensible. We may be inclined to resist the moral implication because we are strongly opposed to e.g. military leaders opting to making direct attacks on civilians when they *feel* justified. Military leaders on opposing sides, may both consider their side the righteous part, when in reality only one, or neither is. Because of this, and in order to avoid attacks based on rampant bias and rationalization, simply outlawing such acts is overall preferable, and will on the whole lead to an overall preferable state of affairs. It is important to have this in mind when gauging our intuitions in the case I will present. It should be read as an ideal moral case, isolated from the legal aspect.

9.4 Moral membership grounds non-combatant liability

I will firstly present the argument for non-combatant liability grounded in moral membership, and then elaborate and discuss the details afterwards. As noted, I will argue that the moral membership of the individual, grounds her liability to attack.

Non-combatant liability argument:¹³⁷

- 1) If elected, politician **P** will make country **A** go to war with country **B**, unprovoked and unjustly.
- 2) “War” necessarily entails the outcome **O** of a “substantial” number of innocent deaths on both sides, destruction of infrastructure and private property.

¹³⁶ E.g. by McMahan 2009: 235.

¹³⁷ I have highlighted the markers in this chapter in **bold** strictly for legibility.

- 3) **P** gets elected by a small majority of citizens, out of which a subset **S** were aware of 1) and 2).
- 4) Therefore, **P** makes **A** go to war with **B**, unprovoked and unjustly, and brings about the harms described in 2).
- 5) Assume each individual of subset **S**, were a likely part of the causally efficacious set of votes that secured **P**'s win, which brought about **O** (and each individual of **S** were aware of this).
- 6) Each individual in subset **S** fulfils the rest of the conditions for moral membership (section 8.4), and are thus morally responsible for **A** going to war with **B**, and the outcome **O**.
- 7) If someone is morally responsible for **O**, they are liable to defensive killing in certain circumstances.
- 8) Therefore, **S** are liable to defensive killing, under those circumstances.

I will now highlight and elaborate some of the above. Premise 7) is most crucial, and will thus receive most of the attention in the following.

Regarding premise 1)¹³⁸, in keeping with the assumption that the rules of war are an extension of the rules governing ordinary conduct in society, we can e.g. specify the unjust war as being the forceful and illegitimate annexation of another country's territory. This would then roughly be tantamount to large scale violent theft initiated by the voters and the president of **A**. So if country **A** goes to war with **B** in order to acquire valuable oil rich land, without historical claim, or moral or legal justification (it had always been part of the territory of **B**), that would constitute an unjust war.

Premise 2)¹³⁹ is a specification of the moral significance of war. This does not require elaboration at this point. It simply states that the outcome is indeed harmful.

¹³⁸ 1) If elected, politician **P** will make country **A** go to war with country **B**, unprovoked and unjustly.

¹³⁹ 2) "War" necessarily entails the outcome **O** of a "substantial" number of innocent deaths on both sides, destruction of infrastructure and private property.

Premise 3)¹⁴⁰ specifies that the relevant voters who voted for P (i.e. the only voters who are in the running for moral membership of the unjust war) will satisfy part of the cognitive condition for moral membership, by being aware of the moral significance of the outcome of their candidate winning. This premise is of course quite implausible in the real world scenario. Firstly, decisions to go to war are usually not advertised prior to an election, for various reasons, a significant one being that it may undermine the war effort to give the defending part time to prepare and garner support from the international community. Further, as already highlighted earlier in the project, it is implausible that all citizens who vote for a candidate are aware of how the actual policy outcome will unfold, or that they even have good reasons to trust the politician's election promises. Still, even in the real world, certain candidates will presumably reasonably be considered more likely to engage in such wars, based on their campaign rhetoric, compared to others. So even if citizens may not have beliefs with that degree of specificity, they presumably do have, or at least ought to have, some expectations regarding the likelihood that their preferred candidate will engage in an unjust war.

Premise/intermediate conclusion 4)¹⁴¹ is simply the conclusion that follows from the previous premises.

Premise 5)¹⁴² specifies that the noted set of voters who voted for P help constitute and sustain the project that brings about the outcome in question. This is the contributory condition from the moral membership account, which along with fulfilment of the cognitive condition, the moral relevance of the outcome, a low cost of defection and upstream control, grounds the individual's moral membership, and thus indirect moral responsibility.

¹⁴⁰ 3) P gets elected by a small majority of citizens, out of which a subset S were aware of 1) and 2).

¹⁴¹ 4) Therefore, P makes A go to war with B, unprovoked and unjustly, and brings about the harms described in 2).

¹⁴² 5) Assume each individual of subset S, were a likely part of the causally efficacious set of votes that secured P's win, which brought about O (and each individual of S were aware of this).

6)¹⁴³ Concludes that the individuals of **S** are morally responsible for the policy outcome, based on their moral membership. I.e. they indeed fulfil all the conditions for moral membership (section 8.4). Piecing this together, we know that the individuals of **S**, fulfil condition 1), i.e. they help constitute or sustain the project that brings about **O**. They also fulfil the cognitive condition. Firstly, they were aware that they helped constitute or sustain the project of the unjust war (specified in 5). Secondly, they were also aware of the significance of **O**, because they were aware of that **P** winning entails going to war with **B**, and they were aware of what going to “war” entails. This awareness was specified in premise 3.¹⁴⁴ Further, the outcome is clearly morally relevant, since it involved e.g. innocent deaths. We can also simply assume that defection was relatively costless, and that the individual has upstream control. In light of this, the individual is a moral member of the project that brings about **O**, and the individual is thus morally responsible for **O**.

Premise/conclusion 7)¹⁴⁵ and its entailment 8)¹⁴⁶ are the most controversial elements and will need to be supported in depth with regards to when something should be considered ‘*defensive killing*’ and what the relevant circumstances are. Moral responsibility may ground liability to attack, but presumably only if there is some further justification involved, at least if we are going to be able to ground accountability beyond mere resentment. Recall from chapter 2, on the assumed *ledger view* (section 2.3), *being* morally responsible, and *holding* responsible came apart. We may conclude that certain voters are morally responsible, because they fulfil the conditions for moral membership. But if we are supposed to *hold* them accountable to a significant extent, we need some justification for this. Specifically, a scenario needs to be specified, where targeting these civilians is

¹⁴³ 6) Each individual in subset **S** fulfils the rest of the conditions for moral membership, and are thus morally responsible for **A** going to war with **B**, and the outcome **O**.

¹⁴⁴ And as noted earlier, we simply assume that the citizen is always sensitive to moral considerations.

¹⁴⁵ 7) If someone is morally responsible for **O**, they are liable to defensive killing in certain circumstances.

¹⁴⁶ 8) Therefore, **S** are liable to defensive killing, under certain circumstances.

justifiable, given the stakes, in light of their moral responsibility for the war. Clearly, given the stipulations so far, it would not seem to follow from the specified premises, that it would always be permissible to kill these voters. There at the very least needs to be *some* further justification in play. That justification may e.g. be at the heart of what should then be considered “defensive killing”. Liability to defensive killing, is killing which “might be necessary to prevent or correct the wrongs involved in the war”, as Jeff McMahan puts it (2009: 218). I will now specify a type of scenario where defensive killing presumably is justified in light of the individual’s moral membership. However, devising such a scenario poses certain challenges. Perhaps the most serious problem is that the relevant group of moral members, may be practically indistinguishable from those who are not moral members. Indeed, part of this may indeed be at the core of our intuitions that civilians are never valid targets in war. As Igor Primoratz writes: *“if civilians supporting an unjust war, by and large, mustn’t be deliberately targeted, that has nothing to do with their rights, their status, their immunity. That is merely a consequence of the fact that bullets, shells, and bombs, by and large, can’t seek them out while staying clear of other, truly innocent civilians within the same population.”* (2002: 241). If this is true, i.e. if part of our intuitions of the non-liability of civilians is due to the practical problem of actually singling them out, then the relevant scenario needs to accommodate this concern. Another challenge, is that it may be quite difficult to accept that the civilians in question indeed were aware of the consequences of their preferred candidate winning, even though it is stipulated. Therefore, the scenario needs to be specified in a way where this is clear. Further, the actual stakes need to be significantly high, to justify targeting civilians.

With the above considerations in mind, we can construct a case where the voters are liable to attack. To support the conclusion of the argument above, we can imagine the following case which fits with the argument above, inspired by a case of from Jeff McMahan (2009:220):

Pre-emptive strike:

Country **A** invades country **B**, unjustly. Country **A** is led by president **P**, a democratically elected megalomaniac, who campaigned on a credible base of restoring the country to former glory by annexing a large inhabited region of country **B**. **P** ran opposed to **P'**, who campaigned on exactly the same base, with the exception that **P'** was adamantly against such an invasion. Everyone is aware that country **A** had no present or historical claim to this region, and the inhabitants do clearly wish to remain citizens of country **B**. The election of **P** itself, was done electronically, from a phone (or other "smart" device) application, by the inhabitants of **A**. The invasion of this region would be extremely costly, and lead to thousands of military and civilian deaths on **B's** side, due to **A's** overwhelming military superiority. President **P** wins the election, and a large destruction of infrastructure and loss of human life is the result.

Country **B** however has few options left (though these are their *only* options). Fearing this scenario in light of the voting campaign, the national intelligence agency of **B**, has been hard at work at developing a computer virus, with which they have now fully infected every phone of **A**. Further, this virus allows them to keep extremely extensive records of all the values, beliefs and activities of the citizens of **A**, among other things giving them perfect information regarding who voted for **P**, and what they were aware of when voting. An added important benefit is that these phones can be set to overload the battery of the phone, causing it to explode, causing severe harm and potential death to any person carrying it. Weeks after the invasion, the virus is now fully in place and ready to be activated. The leader of **B** is presented with five possible scenarios: Option 1) is to accept the annexation, without further fight. Option 2) is to engage in an uprising against **A**. This will expectedly lead to 100,000 injured and dead, soldiers and civilians of **B**, and only a small number of soldiers from **A**. Option 3) would be to overload a large random number of phones of the citizens of **A**, which would expectedly injure and kill 100,000 random citizens of **A** (both people who voted for **P**, and those who did not). Option 4) is to overload the phones of a large number of citizens of **A** who

voted for **P**, which will lead to the injury and death of 100,000 of those citizens. The last option 5) is the same as 4), but with the exception, that they will specifically make sure to target only those who are also moral members of the project. I.e., those who fulfil the conditions for moral membership, including awareness of the significance of the outcome of **P** winning, and with the awareness that they would be a likely part of the causally efficacious set. No matter which strategy we chose aside from 1), it would be sufficient to repel the invasion. Option 3-5 would all be effective strategies because it would show off **B**'s capability and willingness to defend themselves, and produce unacceptable losses for **A**.

Presumably, most will consider 3-5 preferable to 1-2. Further, 4 seems preferable to 3, and most will also consider 5 preferable to 4. This suggests two things. Firstly, it suggests that our pre-theoretic intuitions actually do seem to suggest that there is something clearly incriminating about voting for the evil candidate/policy. Secondly, it suggests that that thing seems to be moral membership. The reason we prefer 5 above all else, is that this is the scenario that most clearly singles out the deeper reasons for why we consider them liable to attack, i.e. the option that is supported by the theory of moral membership. Our theory-dependent intuitions trump our pre-theoretic intuitions. This supports the validity of the moral membership account. Specifically, the fact that certain people voted for the evil candidate and that candidate's policies, and the fact that they made a contribution to that project, in full awareness of the consequences of those policies, holds most of the explanatory force with regards to our intuitions of moral responsibility. I.e. those voters' moral membership in the project of electing the candidate, *and* that project bringing about the policy outcome, seems to be the best explanation for their moral incrimination.

So the argument is complete. Those who are morally responsible via their moral membership are seemingly more clearly liable to defensive killing, and even if there may be other intuitions affecting our judgments here, it seems that moral responsibility – holding everything else equal – pushes our intuitions towards option 5. This suggests that the moral responsibility of those citizens is at the

heart of our intuition of their liability to attack. Further, we would intuitively still prefer option 5 even if we added more victims to that option, but kept victims in the rest of the options fixed. If so, this again supports that moral membership, or something like it, plays a clear role in our intuitions of accountability. Obviously, the moral membership account may still require refinement. Perhaps there is some hitherto unestablished theory of moral responsibility, which to some extent overlaps with the moral membership account, which would lead to the same intuition in this case. In that case, there may be some option 6, which would be preferable. However, as it stands now, the moral membership account seems to be the best explanation for the citizens' liability in the above case.¹⁴⁷

9.4.1 Conclusion

This project concerned one overall question: are individual citizens morally responsible for policy outcomes? In the introduction of this project, I noted a seeming tension between the idea that policy outcomes are representative of the will of the people, as well as a clear reluctance to hold our fellow citizens morally responsible for their voting habits, and by possible extension, policy outcomes. I have now resolved some of this tension. Indeed, I have shown how citizens in principle can be morally responsible for policy outcomes. If the citizens fulfil the conditions for moral membership, then it is appropriate to relinquish our reluctance to hold ourselves and others morally responsible. I.e. it may be appropriate to at least take on the relevant reactive attitude. Further, even if it

¹⁴⁷ A thing to note here, is that we may still have strong intuitions that some sixth option which is not incorporated into the overall scenario, would be preferable, which did not rely on moral membership. Specifically, what if there was an option to only attack frontline combatants, i.e. traditional soldiers currently involved in the A's invasion? Even if we assume that soldiers lack voting rights, my intuition is that we would prefer this option to option 5. This may seem like a problem, because the soldiers are not even morally responsible for the outcome, at least in the same sense. So why would our intuitions favour attacking them? I imagine that this may be due to intuitions relating to posing an immediate threat that comes into play instead, which is not present among the moral members in this case. The voters have already performed their action, and are no longer an immediate threat to anyone. The soldiers on the other hand continue to pose a threat, or are at least seen as posing a threat in our collective consciousness.

may often be practically difficult to ascertain whether a given citizen who voted for a given policy is morally responsible, the fact that it can be appropriate to hold citizens morally responsible, to an extent where they seemingly become valid targets in war, should inspire a greater degree of thoughtfulness when deliberating, and acquiring information, regarding who/what to vote for. There is potentially a great deal at stake.

So it seems we have uncovered part of the theoretical underpinnings for why our intuition might be correct, that the German citizen who voted for the NSDAP in 1933 is morally responsible. To the extent that she actually had some awareness of the significance of that vote, compared to the alternative, that voter is clearly morally responsible, as should have been the intuition all along. Though the German federal electoral system is not a simple two candidate runoff vote, the voter would have been a sufficiently likely part of the set of votes, which secured the NSDAP as the de facto winner.¹⁴⁸ As an interdependent part of the voting project, she helped constitute and sustain Hitler's rise to power, and the ultimate consequences thereof. Therefore, she is morally responsible for those consequences.

10 Bibliography

- Abdel-Nour, F. (2003). National Responsibility. *Political Theory*, 31(5), 693-719.
- Ackerman, B., & Fishkin, J. (2004). *Deliberation Day* (1st ed.). New Haven: Yale University Press.
- Adams, R. M. (1976). Motive Utilitarianism. *The Journal of Philosophy*, 467-481.
- Anscombe, E. (2000). *Intention*. Cambridge, Massachusetts: Harvard University Press.
- Aristotle. (2004). *Nicomachean Ethics*. (R. Crisp, Ed.) Cambridge: Cambridge University Press.
- Bazargan, S. (2013). Complicitous liability in war. *Philosophical Studies*, 165(1), 177-195.

¹⁴⁸ Though they gained the largest percentage of votes, they still needed to broker political alliances in order to establish a working majority.

- Beebe, H., Hitchcock, C., & Menzies, P. (2009). *The Oxford Handbook of Causation*. New York: Oxford University Press.
- Bennett, J. (187). Event causation: The counterfactual analysis. *Philosophical Perspectives*, 367-386.
- Bennett, J. (1981). Morality and Consequences. *The Tanner Lectures on human values*, 2, 45-116.
- Blais, A. (2000). *To voter or not to vote?* (1st ed.). Pittsburgh: University of Pittsburgh Press.
- Boniolo, G., & De Anna, G. (2006). The four faces of omission. *Philosophical Explorations*, 9(3), 277-293.
- Bratman, M. (1987). *Intention, plans, and practical reason* (1st ed.). Cambridge: Harvard University Press.
- Brennan, G., & Lomasky, L. (1985). The impartial spectator goes to Washington: Toward a Smithian theory of electoral behavior. *Economics and Philosophy*, 1(2), 189-211.
- Brennan, G., & Lomasky, L. (1989). *Politics and process* (1st ed.). Cambridge: Cambridge University Press.
- Brennan, J. (2009). Tuck on the Rationality of Voting: A Critical Note. *Journal of Ethics and Social Philosophy*, 3, 1-5.
- Broadie, S. (1991). *Ethics with Aristotle*. New York: Oxford University Press.
- Burnham, G., Riyadh, L., Doocy, S., & Roberts, L. (2006). Mortality after the 2003 invasion of Iraq: a cross-sectional cluster sample survey. *Lancet*, 1421-1428.
- Caplan, B. (2006). *The Myth of the Rational Voter: Why Democracies Choose bad Policies*. New Jersey: Princeton University Press.
- Chan, D. (2000). Intention and responsibility in double effect cases. *Ethical theory and moral practice*, 3(4), 405-434.
- Clarke, R., McKenna, M., & Smith, A. M. (2015). *The Nature of Moral Responsibility: New Essays*. New York: Oxford University Press.
- Cushman, F. (2008). Crime and punishment: Distinguishing the roles of causal and intentional analyses in moral judgment. *Cognition*, 108(2), 353-380.

- Cushman, F. (2013). The role of learning in punishment, prosociality, and human uniqueness. *operation and its evolution (Life and mind: Philosophical issues in biology and psychology)*, 333-372.
- Dennett, D. (1984). *Elbow Room*. Oxford: Clarendon Press.
- Downs, A. (1957). *An economic theory of democracy*. New York: University of Chicago Press.
- Duffy, J., & Tavits, M. (2008). Beliefs and voting decisions: A test of the pivotal voter model. *American Journal of Political Science*, 52(3), 603-618.
- Echeñique, J. (2012). *Aristotle's Ethics and Moral Responsibility* (1st ed.). Cambridge: Cambridge University Press.
- Elster, J. (1998). *Deliberative Democracy*. Cambridge: Cambridge University Press.
- Feinberg, J. (1968). Collective Responsibility. *The Journal of Philosophy*, 65(21), 674-688.
- Feinberg, J. (1970). *Doing & deserving; essays in the theory of responsibility* (1st ed.). Princeton: Princeton University Press.
- Fischer, J. (1999). Recent Work on Moral Responsibility. *Ethics*, 110(1), 93-139.
- Fischer, J. (2012). *Deep Control*. Oxford: Oxford University Press.
- Fischer, J. M. (2006). *My way: Essays on moral responsibility*. New York: Oxford University Press.
- Fischer, J. R. (1998). *Responsibility and control*. Cambridge: Cambridge University Press.
- Fischer, J., & Ravizza, M. (1993). *Perspectives on moral responsibility* (1st ed.). Ithaca: Cornell University Press.
- Foot, P. (1978). *Virtues and vices and other essays in moral philosophy*. Berkeley: University of California Press.
- Frankfurt, H. (1969). Alternate possibilities and moral responsibility. *he Journal of Philosophy*, 829-839.
- Frankfurt, H. G. (1971). Freedom of the Will and the Concept of a Person. *The Journal of Philosophy*, 5-20.
- French, P. A. (1979). The Corporation as a Moral Person. *American Philosophical Quarterly*, 16(3), 207-215.
- Gallup.com. (2010). Retrieved from Gallup: <http://www.gallup.com/poll/157067/views-violence.aspx>

- Gelman, A., Silver, N., & Edlin, A. (2012). What is the probability your vote will make a difference? *Economic Inquiry*, 50(2), 321-326.
- Gilbert, M. (1990). Walking Together: A Paradigmatic Social Phenomenon. *Midwest Studies in Philosophy*, 1-14.
- Gilbert, M. (2002). Review: Collective Wrongdoing: Moral and Legal Responses. *Social Theory and Practice*, 28(1), 167-187.
- Glover, J. (1970). *Responsibility*. London: Routledge & Kegan Paul, New York Humanities Press.
- Goldman, A. (1999). Why citizens should vote: A causal responsibility approach. *Social Philosophy and Policy*, 16(2), 201-217.
- Greco, J. (2006). Virtue, luck and the pyrrhonian problematic. *Philosophical Studies*, 130(1), 9-34.
- Haji, I. (2009). *Incompatibilism's Allure: Principle Arguments for Incompatibilism*. Ontario: Broadview.
- Hart, H. (2008). *Punishment and Responsibility*. Oxford: Oxford University Press.
- Held, V. (2002). Group responsibility for ethnic conflict. *The Journal of ethics*, 6(2), 157-178.
- Herman, E., & Chomsky, N. (1988). *Manufacturing consent*. New York: Pantheon Books.
- Hume, D. (2003). *A Treatise of Human Nature*. New York: Dover Publications, Inc.
- Inwagen, P. V. (1989). When Is the Will Free? *Philosophical Perspectives*(3), 399-422.
- Isaacs, T. (2006). Collective moral responsibility and collective intention. *Midwest studies in philosophy*, 30(1), 59-73.
- Jaspers, C. (2000). *The Question of German Guilt*. New York: Fordham University Press.
- Kane, R. (2012). *The Oxford Handbook of Free Will: Second Edition*. New York: Oxford University Press.
- Kant, E. (2011). *Immanuel Kant: Groundwork of the Metaphysics of Morals*. (M. Gregor, & J. Timmermann, Eds.) Cambridge: Cambridge University Press.
- Kruger, J., & Dunning, D. (1999). Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-Assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-34.

- Kutz, C. (2000). *Complicity*. Cambridge: Cambridge University Press.
- Kutz, C. (2000b). Acting Together. *Philosophy and Phenomenological Research*, 61(1), 1-31.
- Kutz, C. (2007). Causeless complicity. *Criminal Law and Philosophy*, 1(3), 289-305.
- Larvor, B. (2010). Frankfurt counter-example defused. *Analysis*, 70(3), 506-508.
- Lazar, S. (2009). Responsibility, Risk, and Killing in Self-Defense. *Ethics*, 119(4), 699-728.
- Levy, N. (2005). The Good, the Baad, and the Blameworthy. *Journal of Ethics and Social Philosophy*, 1(2), 1-16.
- Levy, N. (2011). *Hard Luck: How luck undermines free will & moral responsibility*. Oxford: Oxford University Press.
- Levy, N. (2013). The importance of awareness. *Australasian Journal of Philosophy*, 91(2), 211-229.
- Levy, N. (2014). Consciousness, implicit attitudes and moral responsibility. *Nous*.
- Lewis, D. (1986). *Philosophical Papers II*. Oxford: Oxford University Press.
- Lewis, D. (2000). Causation as Influence. *Journal of Philosophy*, 97(4), 182-197.
- Mackie, G. (2011). An examination of the Expressive Theory of Voting. *unpublished*, 1-40.
- Mackie, J. (1965). Causes and conditions. *American philosophical quarterly*, 2(4), 245-264.
- May, L. (1992). *Sharing responsibility*. Chicago: University of Chicago Press.
- McIntyre, A. (1994). Compatibilists could have done otherwise: Responsibility and negative agency. *The Philosophical Review*, 453-488.
- McKenna, M. (2012). *Conversation and Responsibility*. New York: Oxford University Press.
- McKenna, M., & Russel, P. (2008). *Free will and reactive attitudes*. Farnham: Ashgate.
- McMahan, J. (2009). *Killing in war*. Oxford: Clarendon Press.
- Mellema, G. (1987). On Risk Taking and Moral Responsibility. *Ethics*, 1-11.
- Mellema, G. (2006). Collective responsibility and qualifying actions. *Midwest Studies in Philosophy*, 168-175.
- Miller, S. (2001). Collective Responsibility. *Public Affairs Quarterly*, 15(1), 65-82.
- Moore, M. (2009). *Causation and Responsibility*. Oxford: Oxford University Press.

- Moore, M. S. (1999). Causation and Responsibility. *Social Philosophy & Policy Foundation*, 1-51.
- Nagel, T. (1979). *Mortal questions*. Cambridge: Cambridge University Press.
- Narveson, J. (2002). Collective responsibility. *The Journal of Philosophy*, 6, 179-198.
- Nichols, S., & Knobe, J. (2008). *Experimental philosophy*. Oxford: Oxford University Press.
- Nolt, J. (2011). How Harmful Are the Average American's Greenhouse Gas Emissions? *Ethics, Policy and Environment*, 3-10.
- Nottelmann, N. (2007). *Blameworthy belief*. Dordrecht: Springer.
- Oakley, J., & Cocking, D. (1994). Consequentialism, moral responsibility, and the intention/foresight distinction. *Utilitas*, 6(2), 201-216.
- Olick, J. (2003). The Guilt of Nations? *Ethics & International Affairs*, 17(2), 109-117.
- Parfit, D. (1984). *Reasons and persons*. Oxford: Clarendon Press.
- Pasternak, A. (2011). The collective responsibility of democratic publics. *Canadian Journal of Philosophy*, 41(1), 99-123.
- Pasternak, A. (2012). Limiting States' Corporate Responsibility. *Journal of Political Philosophy*, 316-381.
- Pereboom, D. (2001). *Living Without Free Will*. Cambridge: Cambridge University Press.
- Pettit, P. (2007). Responsibility Incorporated. *Ethics*, 117(2), 171-201.
- Pettit, P., & Schweikard, D. (2006). Joint Action and Group Agents. *Philosophy of the Social Sciences*, 18-39.
- Pew Research Center. (2003, January 30). Retrieved from <http://www.people-press.org/>: <http://www.people-press.org/2003/01/30/public-struggles-with-possible-war-in-iraq/>
- Pew Research Center. (2003, 2 20). <http://www.people-press.org/>. Retrieved from <http://www.people-press.org/2003/02/20/us-needs-more-international-backing/>
- Pew Research Center. (2008, March 19). Retrieved from <http://www.pewresearch.org/>:

<http://www.pewresearch.org/2008/03/19/public-attitudes-toward-the-war-in-iraq-20032008/>

- Primoratz, I. (2002). Michael Walzer's just war theory: some issues of responsibility. *Ethical Theory and Moral Practice*, 5(2), 221-243.
- Primoratz, I. (2005). Civilian Immunity in War" The Philosophical Forum. *The Philosophical Forum*, 36, 41-58.
- Primoratz, I. (2011). Civilian immunity, supreme emergency, and moral disaster. *The Journal of ethics*, 15(4), 371-386.
- Pryor, J. (2001). Highlights of recent epistemology. *The British Journal for the Philosophy of Science*, 52(1), 95-124.
- Radcliffe-Brown, A. R. (1940). On Social Structure. *The Journal of the Royal Anthropological Institute of Great Britain and Ireland*, 70(1), 1-12.
- Representation*. (n.d.). New York: Atherton Press.
- Riker, W., & Ordeshook, P. (1968). A Theory of the Calculus of Voting. *American political science review*, 62(1), 25-42.
- Rosen, G. (2002). Culpability and ignorance. *Proceedings of the Aristotelian Society* 103, 1, 61-84.
- Sartorio, C. (2001). Causation and Responsibility. *Philosophy Compass*, 5, 749-765.
- Sartorio, C. (2004). How to be Responsible for Something Without Causing it. *Philosophical Perspectives*, 315-336.
- Sartorio, C. (2009). Causation and Ethics. In H. Beebe, C. Hitchcock, & P. Menzies, *The Oxford Handbook of Causation* (p. Ch 26). New York: Oxford University Press.
- Sartorio, C. (2009). Omissions and Causalism. *Noûs*, 43(3), 513-530.
- Scanlon, T. (2008). *Moral dimensions*. Cambridge Massachusetts: Harvard University Press.
- Schoeman, F. (1988). *New Essays in Moral Psychology*. Cambridge: Cambridge University Press.
- Searle, J. (1990). Collective Intentions and Actions. In P. R. Cohen, J. Morgan, & M. Pollack, *Intentions in Communication* (pp. 401-415). MIT Press.
- Sher, G. (2006). Out of Control. *Ethics*, 116(2), 285-301.
- Sher, G. (2009). *Who knew?* Oxford: Oxford University Press.

- Shoemaker, D. (2011). Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility. *Ethics*, 602-632.
- Simmons, A. (1999). Justification and Legitimacy. *Ethics*, 10(2), 130-153.
- Sinnott-Armstrong, W. (2008). *The Cognitive Science of Morality: Intuition and Diversity*. Cambridge, Massachusetts: MIT Press.
- Sinnott-Armstrong, W., & Howarth, R. (2005). *Perspectives on climate change*. Amsterdam: Elsevier JAI.
- Smart, J. (1961). Free-Will, Praise and Blame. *Mind*, 70(279), 291-306.
- Smiley, M. (1992). *Moral Responsibility and the Boundaries of Community: Power and Accountability from a Pragmatic Point of View*. Chicago: The University of Chicago Press.
- Smith, H. (1983). Culpable ignorance. *The Philosophical Review*, 92(4), 543-571.
- Stilz, A. (2011). Collective responsibility and the state. *Journal of Political Philosophy*, 19(2), 190-208.
- Stone, D. A. (1989). Causal Stories and the Formation of Policy Agendas. *Political Science Quarterly*, 281-300.
- Strawson, P. (1962). *Freedom and resentment*. Oxford: Oxford University Press.
- Strawson, P. (1974/2008). *Freedom and resentment, and other essays*. Routledge.
- Sullivan, G. R. (1993). CAUSE AND THE CONTEMPORANEITY OF ACTUS REUS AND MENS REA. *Cambridge Law Journal*, 52(3), 487-500.
- Sverdlik, S. (1987). Collective Responsibility. *Philosophical Studies*, 51, 61-76.
- Sverdlik, S. (1993). Pure negligence. *American Philosophical Quarterly*, 137-149.
- Thomson, J. (2003). Causation: Omissions. *Philosophy and Phenomenological Research*, 81-103.
- Todd, P., & Tognazzini, N. (2008). A Problem for Guidance Control. *The Philosophical Quarterly*, 685-692.
- Tognazzini, N. A. (2013). Responsibility. In H. LaFollette, *The International Encyclopedia of Ethics* (pp. (entry 499) 1-11). Blackwell Publishing Ltd.
- Tuck, R. (2008). *Free Riding*. Cambridge Massachusetts: Harvard University Press.
- Tuomela, R. (2006). Joint Intention, We-Mode and I-Mode. *Midwest Studies in Philosophy*, 35-58.
- Tuomela, R. (2010). Cooperation as Joint Action. *Analyse & Kritik*, 65-86.

- Tuomela, R., & Miller, K. (1988). We-Intentions. *Philosophical Studies*(53), 367-389.
- Vallentyne, P. (2008). Brute luck and responsibility. *Politics, Philosophy & Economics*, 7(1), 57-80.
- Vargas, M. (2005). The trouble with tracing. *Midwest Studies in Philosophy*, 29(1), 269-291.
- Vincent, N., Poel I, & Hoven, J. (2011). *Moral Responsibility: Beyond Free Will and Determinism*. Dordrecht: Springer.
- Walzer, M. (1973). Political action: The problem of dirty hands. *Philosophy & public affairs*, 160-180.
- Walzer, M. (1977). *Just and unjust wars*. New York: Basic Books.
- Watson, G. (1982). *Free will*. Oxford: Oxford University Press.
- Watson, G. (1987). Responsibility and the Limits of Evil. In F. Schoeman, *Responsibility, Character and the Emotions* (pp. 256-286). Cambridge: Cambridge University Press.
- Watson, G. (1996). Two Faces of Responsibility. *Philosophical Topics*, 24(2), 227-248.
- Watson, G. (2004). *Agency and answerability*. Oxford: Clarendon Press.
- Williams, B., & Nagel, T. (1976). Moral Luck. *Aristotelian Society Supplementary Volume*, 115-151.
- Wolf, S. (1980). Assymetrical Freedom. *The Journal of Philosophy*, 77(3), 151-166.
- Wolf, S. (1981). The Importance of Free Will. *Mind*, 386-405.
- Wolf, S. (1990). *Freedom Within Reason*. Oxford: Oxford University Press.
- Zimmerman, M. J. (1988). *An Essay on Moral Responsibility*. Totowa, New Jersey: Rowman and Littlefield.
- Zimmerman, M. J. (2002). Taking luck seriously. *The Journal of Philosophy*, 553-576.
- Zimmerman, M. J. (2008). *Living with Uncertainty: The Moral Significance of Ignorance*. Cambridge: Cambridge University Press.
- Zimmerman, M. J. (2010). Responsibility, Reaction, and Value. *The Journal of Ethics*, 14(2), 103-115.