

Authority, Delegation, and Plans

Ludovica Adamo

The University of Leeds
School of Philosophy, Religion and History of Science

September 2023

To Maggie & Jamie

Acknowledgements

Many people have supported me, one way or the other, throughout this journey.

Firstly, I wish to thank my supervisory team, Gerald Lang, J. Robert G. Williams, and Daniel Elstein. This project would have looked a lot different without your patient guidance and support. Thank you for helping me shape my thoughts and form my views. Thank you, Robbie, for our productive brainstorming sessions and the lovely GROUNDS walks. You welcomed me into your team with open arms and I will be forever grateful for that. Thank you, Gerald, for your immense patience and positivity. There wasn't a single time I left your office after a meeting that I didn't feel ten times better than when I entered it. You have the power to make everyone feel at ease, and it is one of the many things I admire about you. This experience would have not been the same without you. Thank you, Daniel, for your caring support and for helping me develop as a philosopher.

I wish to thank ERC and the GROUNDS project for their essential financial support in the past four years. I am also grateful to my examiners, Saba Bazargan-Forward and Jamie Dow, for dedicating their time to my work.

Many of the ideas in this thesis have been inspired by and developed through discussions with others. I am grateful to Thomas Brouwer, Haixin Dang, Joshua Habgood-Coote, Kirk Ludwig and Abraham Roth.

I am forever grateful to my loving and caring partner Jamie. Thank you for always putting a smile on my face and believing in me. I lost myself in this project. Knowing you'll be by my side in the process of healing and finding myself again fills me with hope and gratitude. Thank you for charting all the uncharted territories and stroking all the dogs with me in the past three and a half years.

A huge thank you to the lovely people of Botany House, who have made this experience much more enjoyable than it would have otherwise been. Very special thanks to Maddy Page and Jill Schacher, for the laughs and the many girls' nights. Special thanks to Simon Graf, Sam Mason, and Rodrigo Valencia. I also wish to thank our adoptive Botany House member, Dawn Wang.

Special thanks to Sofia Lombardi, my lovely friend and co-author. Your support throughout this project has been invaluable. Walking the same path as you has been a huge source of comfort and inspiration.

I also wish to thank my St Andrews Flat 4 familia. Thanks to Shubhi, David, and Juan for supporting me, believing in me and always being there for me. You inspire me to be the best version of myself. I will be forever grateful to many other members of the St Andrews crew, including William and Kostas.

I am deeply indebted to my team members at *Let's Phi*, especially Konrad Urban and Joshua Bucheli. The employability work we have done for students and graduates in the past three years has fulfilled me in many ways and created opportunities for personal and professional development. There's still a lot of work to be done to support students and make academia a better place. Thank you for allowing me to be a part of this mission.

Finally, I am grateful for Maggie, Nature, its beautiful animals, and Darin Olien.

Abstract

Understanding how large-scale hierarchical groups, such as corporations and nation-states, act is a theme of the recent literature on joint action (Bratman 2021; Gilbert 2006; Ludwig 2017; Shapiro 2011). But while several accounts analyse the practical mechanisms of large-scale agency, more focus is needed on the normative questions that these mechanisms raise. Shapiro (2011), for instance, notes that states coordinate their subjects via plans to reach common goals. This large-scale co-ordinational activity is necessary to achieve social order. However, from the practical necessity for large-scale, social coordination, and the fact that states can achieve it, it does not follow that it is legitimate for them to do so – for example, the authority’s planning seems to impair personal autonomy. My research develops a theory of political authorities as large-scale hierarchical groups that answers normative questions concerning the legitimacy of political authorities and the relationship between states and their subjects.

This thesis is divided into two parts. Part 1 proposes a three-stage model of practical reasoning that demonstrates how individuals can remain autonomous despite participating in hierarchical group action. Personal autonomy is preserved because authority’s plans help individuals respond to reasons and even enhance their reason-responsiveness capacities. I also note that there seems to be a more general tension between autonomy and shared agency that stems from the fact that plans need to be treated as exclusionary to guide our conduct. The three-stage model of practical reasoning ultimately shows that autonomy and shared agency, that is, being autonomous and acting with others, are compatible.

Part 2 addresses the legitimacy problem. It defends what I call *Delegation Theory of the Legitimacy of Political Authorities*. According to this theory, political legitimacy is based on the authority producing adequate social plans and providing a basic public justification for them. When these two conditions are met, the subjects have a duty to delegate their social

planning powers to the authority. I then argue for normative delegation, which is a non-voluntarist theory of legitimate political authority. This research puts forward an account of political legitimacy understood as a product of large-scale, hierarchical agency, thus contributing to the emerging field of political ontology.

Table of Contents

Acknowledgements	3
Abstract.....	5
Chapter 2 – The Autonomy-Authority Problem Revisited:	17
Conceptual Preliminaries	17
2.0 Introduction.....	17
2.1 Wolff’s Autonomy-Authority Problem	20
2.1.1 <i>Challenging Authority</i>	24
2.1.2 <i>A Different Take on Autonomy</i>	28
2.1.3 <i>The Autonomy-Authority Problem Reformulated</i>	30
2.2 Three theories of the legitimacy of political authority.....	31
2.2.1 <i>Consent theories of legitimacy</i>	31
2.2.2 <i>Membership-based accounts of legitimacy</i>	46
2.2.3 <i>The Instrumentalist View</i>	53
Chapter 3 - Individual Practical Reasoning, Plans and Autonomy.....	61
3.0 Introduction.....	61
3.1 Intentions and Plans	64
3.2 Explanatory models of planning and plan-related reasons	72
3.3 Exclusionary reasons, autonomy and personal identity.....	79
3.4 Summary	87
Chapter 4 - Shared Practical Reasoning, Autonomy and Shared Plans	89
4.0 Introduction.....	89
4.1 Joint Activity and Shared Plans	91
4.2 An explanatory model of shared practical reasoning.....	100
4.3 Autonomy and the exclusionary nature of shared plans	105
4.4 Summary	117

Chapter 5 - Hierarchical Shared Agency & Hierarchical Groups:	119
How Authorities and their Subjects Do Things Together	119
5.0 Introduction	119
5.1 Hierarchical Shared Agency	120
5.1.1 <i>Transmission of Reasons</i>	120
5.1.2 <i>Hierarchical Shared Agency and Group Hierarchy</i>	142
5.1.3 <i>Nation-states as hierarchical groups</i>	154
5.2 An account of hierarchical shared agency for nation-states	159
5.3 Individual Autonomy and Authoritative Plans	173
5.4 Summary	182
Chapter 6 - The Delegation Theory (Part 1)	184
6.0 Introduction	184
6.1 Authoritative plans	185
6.2 Public Justification on The Part of The Authority	192
6.3 Summary	202
Chapter 7 - The Delegation Theory (Part 2): A Defence of Normative Delegation	204
7.0 Introduction	204
7.1 Delegation	205
7.2 Delegation vs. Consent	212
7.3 Voluntary vs. Normative Delegation	216
7.3.1 <i>The two conditions of legitimacy and delegation</i>	228
7.4 Objections to normative delegation	230
Chapter 8 – Conclusion	238
Bibliography	246

Chapter 1 – Introduction

“But the social order is a sacred right, which is the basis of all other rights. Nevertheless, this right does not come from nature, and must therefore be founded on conventions.”

Rousseau, *The Social Contract*, Book I, Chapter I.

“All that makes existence valuable to any one, depends on the enforcement of restraints upon the actions of other people.”

Mill, *On Liberty*, Chapter 1, P. 9.

Greenpeace accuses Shell and BP of greenwashing over renewable energy. Google’s board of directors lays off 6% of the company’s workforce. The UK government institutes a lockdown and orders its citizens to stay home to contain COVID-19. These are actions where the agents involved are members of large-scale, hierarchical groups. Understanding how large-scale hierarchical groups act is a theme of the recent literature on joint action (Bratman 2022; List & Pettit, 2011; Ludwig 2017; Shapiro 2011; Tollefsen, 2015; Tuomela, 2013).

There are various views on large-scale hierarchical groups and their agency, and I am going to briefly canvass some of them. In his recent book, Michael Bratman argues that agents’ capacity for planning underpins different forms of multi-agent, organised action, including the actions and organisation of institutions (Bratman, 2022). Bratman contends that institutions can have intentions and be intentional agents despite not being rational agents in the way that individuals are. Christian List and Philip Pettit (2011), like Bratman, analyse group agents and argue that they are real and irreducible to individual agents and their mental states. According to them, group agents are not mysterious entities that supervene upon the aggregate of their individuals. However, they can exhibit representational and motivational states that, when adequately processed, allow them to act in their environment.

Kirk Ludwig also investigates the agency of institutions, including corporations and nation-states (2017). Ludwig provides us with an influential account of how these large-scale groups are structured and how that structure, which consists of collectively accepted roles and functions, allows the persistence of these groups over time. Institutional agency is, for Ludwig, the agency of the group's members, who are connected via a socially constructed membership relation. Scott Shapiro offers a new take on legal positivism when he argues that the social facts that determine legal facts are facts about plans (2011). He argues for the *Planning Theory of Law*, according to which the exercise of legal authority is an activity that involves social planning. Legal activity, the creation and enforcement of laws, is then an instance of shared agency. Legal officials create and apply the law together, thus acting together and sharing intentions.

Deborah Tollefsen (2015) argues that certain groups should be treated as agents and held morally accountable for their actions. Raimo Tuomela (2013) investigates group thinking, collective intentions, social reasons and so forth. He argues for collective intentions in the we-mode as mental states that are irreducible to individual mental states, despite not being ontologically distinct from individual mental states. Tuomela ascribes these collective mental states to groups. He also makes an important distinction between we-reasons, the group's reasons, and the reasons of individuals.

These accounts brilliantly analyse the practical mechanisms of large-scale agency, that is, the internal structure, organisation, and functioning of groups such as institutions, corporations, and nation-states. While this analysis is central to understanding how members of large-scale groups do things together, more focus is needed on the normative questions that the practical mechanisms of large-scale agency raise. The theory of action literature has started making some progress in this area by answering some normative questions about group agency.

Deborah Tollefsen, as previously stated, has investigated ascriptions of legal and moral responsibility to groups that people make in everyday speech and argues that some groups should be held morally accountable for their actions (Tollefsen, 2015). Similarly, Jennifer Lackey's work on the epistemology of groups and group lies can be used as a foundation to engage in debates on group responsibility and group liability (Lackey, 2020). For instance, if we can say that a group is lying, we can hold it (or some of its members) responsible for the spreading of misinformation and the harm that misinformation caused.

Furthermore, other scholars have answered other normative questions that group agency raises. Saba Bazargan-Forward, for instance, has developed a theory of individual moral accountability in the social world. This account can help us understand how to hold individuals accountable for their actions when they are acting with multiple agents in a way that results in harm (Bazargan-Forward, 2022). Moreover, Stephanie Collins' work on group duties delves deeper into the assumption that groups can have moral responsibility (Collins, 2019). Her account sheds light on how membership in a group generates duties for its members.

Margaret Gilbert also addresses the problem of political obligation, thus answering some normative questions related to political societies and nation-states as large-scale groups (Gilbert, 2006). She argues that it follows from being a member of a political society that one has an obligation to support the political institutions of that society. Gilbert applies her renowned account of joint commitment to political societies as group agents whose members acquire obligations towards the group by jointly committing to something as a body. Her work is highly influential in political philosophy and has advanced our understanding of political obligations through action theory.

However, other normative questions need to be addressed. Let us go back to Shapiro's *Planning Theory of Law*.

Shapiro's theory of how authorities plan for their subjects and how the subjects intend to obey authoritative plans gives rise to what he calls 'rational authority' (Shapiro, 2014, p. 267). When authorities create plans that their subjects follow, the authority's and the subject's intentions are intertwined and governed by norms of practical rationality. If individuals submitted to the authority's orders and then did otherwise, they would be acting irrationally, as they would be going against their intention to obey the authority. If a state has rational authority, however, it does not mean that it is reasonable or advisable to obey it. The state's plans might be misinformed, misguided or immoral and individuals would not have a duty to obey them. It might be rational for them to follow the authority if they intend to do so, but that does not mean that they should.

Shapiro rightly notes that states coordinate their subjects via plans to reach common goals. This large-scale co-ordinational activity is necessary to achieve social order. However, from the practical necessity for large-scale, social coordination, and the fact that states can achieve it, it does not follow that it is legitimate for them to do so - for example, the authority's planning seems to impair personal autonomy.

In this thesis, I answer normative questions concerning the legitimacy of political authorities and the relationship between states and their subjects. Building on Shapiro's Planning Theory of Law, I argue that political authorities can be legitimate when their plans meet certain conditions. When the authority's plans are adequate, the subjects have a duty to obey them. It is not just rational for the subjects to follow political authorities that plan in a certain way, it is justified and reasonable for them to do so. I also argue that legitimate political authority and individual autonomy can coexist and be compatible. Ultimately, the thesis proposes an account of political legitimacy understood as a product of large-scale, hierarchical agency.

Aside from this chapter, which serve as an introduction, the thesis is divided into two parts.

Part 1 investigates the relationship between authorities and their subjects, starting from Robert Paul Wolff's claim that legitimate authority and autonomy can never coexist (Wolff, 1970). The tension between legitimate authority and autonomy stems from the fact that authoritative plans ought to be treated as exclusionary. The subjects are not supposed to deliberate on the merit of authoritative plans. They simply must follow them. A more detailed analysis of planning agency, however, reveals that any plan, not just plans created by authorities, must be treated as exclusionary to guide agents' conduct. There is then a *prima facie* general tension between autonomy and shared agency that stems from the fact that any plan is exclusionary in nature. Part 1 of the thesis puts forward a three-stage model of practical reasoning that analyses individual planning agency, shared agency (individuals acting with others as equals) and hierarchical shared agency (individuals acting together under an authority). The model shows that individuals can remain autonomous when they participate in shared agency, with or without the presence of the authority. Personal autonomy is preserved in instances of egalitarian or hierarchical shared agency because plans can help individuals respond to reasons and even enhance their reason-responsiveness capacity. Plans enhance agents' reason-responsiveness capacity as they can store reasons for action and transmit them from agent to agent. I will argue that authorities can transmit reasons to their subjects via plans and that the authority's reasons can thus motivate the subjects to act in accordance with authoritative plans. I defend a version of Abraham Roth's entitlement to other agents' reasons account applied to authorities and their plans (Roth, 2017).

Part 2 of the thesis addresses the legitimacy problem. It presents the *Delegation Theory*, according to which political authorities are legitimate when they produce adequate social plans for their subjects and provide a public justification for their plans. When these two conditions

are satisfied, the subjects have a duty to delegate their social planning powers to the authority. I then defend a non-voluntarist theory of legitimate political authority that builds on David Estlund's notion of normative consent (Estlund, 2008).

I now provide a brief overview of the thesis chapters. Part 1 of the thesis includes Chapters 2, 3, 4 and 5. Part 2 includes Chapters 6, 7 and 8.

Chapter 2 examines Wolff's autonomy-authority problem. There is a tension between deciding for yourself what to do and following authority's commands or, in general, following another agent. This tension stems from having to treat authoritative plans and directives as exclusionary (Raz, 2006). When we treat plans as exclusionary, those plans stop our deliberation and filter out some courses of action that become unavailable to us. It seems as though we cannot decide for ourselves when we follow the authority and treat its plans as exclusionary. Deciding for ourselves, however, is important. So, how to solve this tension? I embrace the Razian view of authority according to which authority gives its subjects exclusionary reasons. I distance myself from Wolff's autarkic view of autonomy, which sees autonomy as a duty to give oneself moral commands. I propose a more neutral account of autonomy as the exercise of the capacity to decide for ourselves and respond to reasons. I suggest, however, that the tension between deciding for ourselves and following the authority remains. To solve this tension, I follow Raz's theory of legitimate authority, according to which authorities can help us respond to reasons. I do this with some modifications.

In Chapter 3, I explain that the tension between legitimate authority and autonomy is part of a more general tension between autonomy and shared agency, that is, between autonomy and acting with others, irrespective of authority relations within a group of agents. This tension arises because all plans must not be (at least initially) reconsidered and treated as exclusionary to guide our conduct. Plans we create with others have constraints that are the product of

different agents' practical reasoning and, in following them, we might be following someone else's reasons, preferences or choices. It is not immediately clear that our autonomy is safe when we act with others if we treat shared plans as exclusionary. To show how autonomy is not compromised in cases of shared agency with or without authority, I propose a three-stage model of practical reasoning. Chapter 3 delves deeper into the first stage of the model: individual practical reasoning. It highlights how treating plans as exclusionary facilitates our decision-making processes and helps us respond to reasons by reconnecting us with the reasons for action that went into the creation of our plans made in the past.

Chapter 4 analyses the second stage of the model, egalitarian shared agency. Egalitarian shared agency is about agents planning and acting with others without hierarchy among them. In this chapter, I argue that shared plans, the plans of action we create and execute with other agents, do not impair personal autonomy. Shared plans can help us respond to our reasons when coordinating and organising our conduct with others is essential to achieve our goals.

In Chapter 5, I analyse the last stage of the practical reasoning model, which deals with hierarchical shared agency. Hierarchical shared agency encompasses those cases where agents do things together under the authority's direction. I explain how plans created by an authority for its subjects can help us to respond to reasons, just like any other plan would do. Moreover, authoritative plans enhance our reason-responsiveness capacity when a transmission of reasons occurs from the authority to us via those plans. In this chapter, I explain my account of transmission of reasons, I compare this account to similar approaches in the literature from other authors and defend it from criticism. This chapter concludes the analysis of my three-stage model of practical reasoning. Ultimately, the model gives us a picture of autonomy that is not based on total self-reliance and is instead compatible with acting together and delegating tasks to others, authorities included.

Chapter 6 is the start of Part 2 of this thesis. After showing that autonomy and legitimate authority can be compatible, I turn to the legitimacy question. Political authorities are legitimate when they produce adequate plans for their subjects and when they publicly justify those plans. The desideratum for adequate plans is that they need to be sensitive to solving societal problems. Authorities also need to give the subjects a chance to evaluate the adequacy of their plans if needed via public reasoning. This gives the subjects the possibility to decide whether they should follow the authority or do otherwise in circumstances when they might have more stringent reasons to do otherwise. This chapter goes into detail on these two conditions of legitimacy authorities need to meet.

In Chapter 7, I argue that when those two conditions of legitimacy are met by the authority, its subjects ought to delegate their social planning powers to the authority. This idea of a duty to delegate on the part of the subjects (normative delegation) is inspired by urgent task theories. Societal problems are very pressing to solve, and the state is in a better position than us to solve them. When this is the case, agents have an obligation to delegate planning to the authority that can solve them. And we said that those authorities are legitimate when they plan in a certain way. So, agents do not have a duty to delegate to *any* authority. Chapter 7 also addresses two objections that might be raised against normative delegation.

In the Conclusion (Chapter 8), I reiterate my defence of the compatibility of autonomy and authority, as well as my defence of normative delegation. I also present some interesting applications of my theory of legitimate political authority for further research.

Chapter 2 – The Autonomy-Authority Problem Revisited:

Conceptual Preliminaries

2.0 Introduction

Do agents lose their autonomy when they are subjected to authority? How does abiding by the law affect individual practical reasoning? These questions have been extensively discussed in the debate on legitimate political authority (Edmundson, 1998; Green, 1989; Hart 1961; Raz, 1986, 1990; Shapiro, 2004; Simmons, 2001; Weber, 1970). Philosophical anarchists claim that the state's authority negatively impacts agents' capacity for acting autonomously to the point where Robert Paul Wolff (1970) argues that there is an inherent incompatibility between the two. Autonomy is the exercise of individual self-governance, while authority consists in a right to rule exercised through commands (Wolff, 1970). If we are to obey the authority and its commands, we cannot exercise our autonomy and give ourselves orders. The autonomy of a rational agent and the authority of the state appear to be on a collision course.

Some responses to Wolff accept the incompatibility of these understandings of authority and autonomy (Shapiro, 2004). Other strategies try to reconcile them and show that they can coexist (Raz, 1986, 2006). Similarly to Raz, I argue that we can reconcile autonomy and authority. In doing so, we need to understand how exercising our capacity to make decisions and respond to reasons is compatible with the peremptory or exclusionary nature of authoritative directives.

According to the view on authority I subscribe to,¹ legitimate authority has a right to rule that it primarily exercises by issuing laws and directives. These are thought to give the subjects

¹ Wolff (1970) also seems to have in mind this understanding of authority in his argument for philosophical anarchism that I reconstruct in section 1.

exclusionary reasons for action. When the subjects follow these reasons and act upon them, their deliberative processes are blocked, and they do what they are ordered because the authority said so. I define autonomy as the exercise of the capacity to plan one's conduct (i.e., decide for oneself) and respond to reasons.² Planning one's conduct amounts to settling the matter on how to act by forming intentions concerning present or future actions and the means to achieve those actions. Agents are responsive to reasons when they weigh reasons for and against actions and choose the one (they believe) that best conforms to the balance of reasons.³ In short, we can say that autonomous agents decide what to do in a way that is responsive to reasons. When these two conditions, deciding what to do and responding to reasons, are satisfied, an individual is autonomous.

It appears that the exclusionary nature of authoritative directives is not compatible with individual autonomy. When they are autonomous, individuals decide for themselves what to do and act upon reasons; when they obey the authority, they do not deliberate, and their practical reasoning is influenced by the authority. Raz (1986) would argue that these two elements can be reconciled when the authority issues directives that are based on the subject's reasons. If the subjects can conform to their balance of reasons by following directives that are based on them, the tension between the exclusionary reasons that stem from authoritative directives and the subjects' deliberation seems to disappear. Authoritative directives that help the subjects respond to reasons (Raz's solution to Wolff's paradox) are a step in the right direction in reconciling autonomy and authority. However, I believe we can go a bit further. What about those cases where the authority does not base its directives on the subject's reasons? Is that authority illegitimate? Not necessarily. Following authoritative directives, whether they

² This view on autonomy differs from Wolff's autarkic conception of personal autonomy. More details on the two positions are to follow in section 1.

³ Being responsive to reasons in this way is compatible with an agent's beliefs about her reasons being generally false, so long as her actions are in some significant way guided by these (false) beliefs about her reasons.

are based on our reasons or not, can help us fulfil our goals and do what we have the most reason to do.

I, then, aim to solve R. P. Wolff's (1970) paradox by showing that the incompatibility between autonomy and political authority is only apparent. Through the theory of action, namely, through an analysis of how individual practical reasoning is influenced by and intertwines with authority's practical reasoning, I show that authoritative directives can help us respond to reason (even when they are not based on our reasons).

In this chapter, I examine Wolff's paradox of autonomy and authority and analyse how three influential theories of the legitimacy of political authority try to solve its reformulated version. Section 2.1 presents Wolff's version of the paradox and explains why it is important for political philosophers to respond to his anarchist challenge. Sections 2.1.1 and 2.1.2 examine two possible ways to solve the paradox. In section 2.1.1, I look at those theories that weaken the notion of legitimate authority, while in section 2.1.2, I offer a less committal definition of autonomy than Wolff does. Modifying either authority or autonomy (or both) is supposed to render Wolff's paradox null. I reformulate the paradox and I show that the tension between autonomy and authority remains. Section 2.2 deals with three theories of the legitimacy of political authorities and how they take on Wolff's challenge. Section 2.2.1 deals with consent theories of legitimacy; section 2.2.2 analyses membership-based accounts and section 2.2.3 the instrumentalist view of authority. I focus on how these theories would respond to the new 'anarchist' challenge posed by my reformulation of Wolff's paradox. In section 2.3, I summarise the chapter.

2.1 Wolff's Autonomy-Authority Problem

Wolff (1970) claims that individual autonomy and the authority of the state cannot coexist due to their inherent incompatibility. The very idea of an autonomous individual subjected to the state is nonsensical to Wolff. We can reconstruct the argument that supports his conclusion as follows:

P1: Obeying the authority means doing what the authority tells you to do *because* the authority tells you to do it.

P2: Autonomy, as self-legislation, entails having a duty to give yourself moral commands.

P3: It is impossible to do something *because* the authority tells you to do it and to be self-legislating at the same time.

C: Therefore, authority and autonomy cannot coexist.

The notion of authority Wolff has in mind in P1 applies to those practical authorities that purport to direct and influence individuals' behaviour.⁴ Political authorities direct those under their jurisdictions by issuing laws and directives for their subjects to follow. Sometimes this is described as the political authority's right to rule which, might be accompanied by a duty to obey on the part of the subjects (Shapiro, 2004). What is distinctive about this understanding of authority is that when the subjects are presented with authoritative directives, they are not generally supposed to assess them and decide whether to follow them. They seem to have a reason to obey that entails conforming to certain commands, regardless of their evaluation of

⁴ See Adams (2018); Star and Delmas (2011) and Westlund (2018) for more on practical authority.

them. We might also think that authoritative commands ought to be obeyed because they were issued by the authority itself, independently of their specific content.⁵

Regarding the view of autonomy in P2, Wolff, who is liberally following Kant (1996), argues that being autonomous is a moral duty. This duty is grounded on agents' will being free and their ability to engage in practical reasoning regarding how to act. Given that one can reflect on what to do and act upon related intentions and plans, one is under an obligation to "make himself the author of his decisions" (Wolff, 1970, p. 18), that is, an obligation to self-legislate.

Why are authority and autonomy as defined above in tension? As it stands, Wolff's argument seems both obvious and unappealing. It seems obvious because the kind of paradoxical tension he invokes to support his conclusion that personal autonomy and political authority are incompatible is, somewhat, not very illuminating. There is no doubt that abiding by authoritative commands opposes giving yourself (moral) commands. Authoritative commands are given to you by someone else, and (moral) commands you give yourself come from within. They may sometimes be opposite in their content too, for authoritative commands may clash with your moral commands, such as when an army officer orders a soldier to kill the enemy and the soldier takes himself to have the moral imperative not to kill. Understood this way, the tension between autonomy and authority is the tension between running your own life in a certain way and someone else (i.e., the authority) giving you commands, thus running your life on your behalf (to a certain extent at least).

But this does not illuminate the relationship between the autonomous individual and the authority in any significant way. It does not tell us why it is preferable to be in charge of our own life as opposed to following someone else's commands or vice versa. It might be that

⁵ These two features of authoritative directives are what H. L. A. Hart (1982) calls "peremptory" and "content-independent" reasons for action (pp. 253-255). A peremptory reason excludes assessment from the individuals, while a content-independent one ought to be followed because an authority provided it, regardless of its content.

following authoritative commands shapes our life in a better way than running it ourselves does. Moreover, it is not clear that from that kind of tension, it follows that autonomy and authority cannot coexist. We might think that finding a balance between obeying the authority and doing as we please is very well achievable and, perhaps, preferable to the state of conflict that Wolff's anarchism might put us in.

Wolff's account could also strike as unappealing due to what P1 and P2 commit us to if we accept them. On the one hand, the idea of 'blind' obedience promoted by P1, according to which complying with an authority entails doing what we are told because the authority said so is, at least *prima facie*, wrong. Why would we (accept to) do as someone told without first reflecting on the matter and establishing for ourselves if that ought to be done? We might be ordered to do something immoral or irrational. Why, then, would we (accept to) do something without first ascertaining that what we are ordered to do is morally right or rational? Wolff's notion of authority appears to be very strong and we might want to seek one that does not involve blind obedience to the people in authority.

On the other hand, P2's conception of autonomy is also very strong. Wolff claims that the duty to give ourselves moral commands is something we should always strive to achieve. We are autonomous when we fulfil this duty to give ourselves moral imperatives.⁶ Does this mean that this duty cannot be overridden? There may be cases when following someone else's orders trumps our duty to give ourselves moral commands. If I am on a sinking ship, following the captain's order regarding what to do to save myself and other passengers might well override the duty I have to give myself commands. Can I not decide to follow orders, where that decision counts as an expression of my autonomy? Or do I (temporarily) lose my autonomy

⁶ It is worth mentioning that giving myself moral commands is not a sufficient condition for the adequacy of those commands. I cannot bootstrap myself into acting in a morally proper way. I might command myself to treat everyone's property as mine. This may be a self-imposed moral duty, but it is clearly morally deficient. Wolff does not seem to analyse this issue adequately.

when I follow commands given by someone else? Wolff does not fully explain the relationship between the duty to be autonomous and other duties one can have. And we might think that the autarkic ideal of an agent constantly giving herself moral commands is unsustainable, since agents are interconnected and depend on one another in social settings.

For these reasons, we might think that Wolff's account does not shed light on the delicate balance between autonomy and authority. Nevertheless, Wolff is not merely pointing to a tension between running your own life and following authoritative commands. Rather, he is pointing at the contrast between deciding for yourself how to act and following another. This notion of autonomy as deciding for yourself how to act entails being the final arbiter of your decisions and actions. This is connected to taking responsibility for your actions. Agents can, in paradigmatic cases at least, determine what to do, where this involves gaining knowledge of a situation, weighing reasons for or against certain actions and, ultimately, settling on a course of action. It is via this process of deciding for themselves that agents take responsibility for their actions, and it is because of this process that others consider them responsible for their actions. Deciding for yourself is, then, tied to moral responsibility. When we understand autonomy this way, the tension between autonomy and authority becomes the tension between deciding for yourself, thus taking responsibility for your actions, and not being able to decide for yourself, because the authority provides commands that hijack your deliberative processes. How can we take responsibility for our actions if we cannot decide for ourselves what to do?

We can see how P1 and P2 might clash in this sense. If one considered what the authority ordered and ultimately decided for oneself what to do, the authority would be useless, for its commands could be easily discarded. Authorities are in power for specific reasons, and they need to find ways to enforce their laws and seek compliance from the subjects. A way to induce compliance is by not permitting the subjects to deliberate on what to do when presented with authoritative commands (P1). In this sense, when one is subjected to authority, one cannot

decide for oneself. P2, though, tells us that we ought to decide for ourselves. If we do not do that, we cannot take responsibility for our actions and not taking responsibility is morally bad. So, deciding for oneself is, in a way, incompatible with authority.

We might think that either P1 or P2 is false, and we might try to solve this tension by rejecting one or more of the premises, but we can now see the appeal of P1 and P2 in the sense I described. In what follows, I argue that P1's understanding of authority is indispensable for the proper functioning of political societies, and I reformulate the problem by altering P2. Even on a different, more neutral, understanding of autonomy, the tension with authority remains.

2.1.1 Challenging Authority

I now discuss those accounts that modify what I call the strong view (SV) of political authority. One way to challenge Wolff's version of the autonomy-authority problem and its anarchist conclusion is by rejecting P1's notion of authority. I do not see this as a viable way to solve his paradox.

P1 states that obeying an authority means doing what the authority tells you to do *because* the authority tells you to do it. I call this the *strong view* on (political) authority (SV, hereafter). SV is based on Raz's (2006) account of authority, which inquires into the nature of authoritative reasons. According to Raz, authoritative directives provide exclusionary reasons for action (1986; 2006), that is, second-order reasons that are meant to be substituted for the subjects' first-order reasons. Exclusionary reasons are reasons to act according to authoritative directives, which exclude the subjects' reasons that are contrary to the content of the law. They also exclude the subjects' deliberation (and first-order reasons) by providing them with second-order reasons to do what the authority tells them to do (Ehrenberg, 2011). For Raz, authoritative directives provide reasons to do as the authority says, which are accompanied, in the case of

legitimate authority, by correspondent exclusionary reasons (reasons not to do something that goes against the content of the directive) (Raz, 2006, p. 1002). The combination of these two kinds of reasons is called a ‘pre-emptive’ reason. Exclusionary reasons act as a filter in the subjects’ practical reasoning: they filter out reasons for action that are in contrast with the directive, thus actively leading the subjects to do as the authority says.

Similarly, H. L. A. Hart calls ‘peremptory’ those reasons that exclude assessment from the individuals by “cutting off deliberation, debate or argument” (1982, p. 253). Hart argues that authorities provide these peremptory reasons that do not function as normal, first-order reasons the subjects have. First-order reasons are reasons to do something that can be weighed against a variety of other first-order reasons. Peremptory reasons are not to be weighed against other first-order reasons, and they are not even the strongest (weightiest) first-order reasons that would automatically outweigh any other reasons. They are reasons which exclude the subjects’ deliberation so that the subjects conform to the directive issued by the authority.⁷

What seems problematic about SV is obedience based on exclusionary reasons. These reasons limit the subject’s deliberation by filtering out reasons that are inconsistent with the authoritative directives so that directives can be treated as conclusive reasons for action (Ehrenberg, 2011, p. 886). And this affects individuals’ capacity to decide for themselves what to do, as some options they can choose from in deciding how to act are ‘eliminated’ by the exclusionary power of authoritative directives. This seems, on the face of it, controversial. Rational agents weigh different courses of action and choose the ones that allow them to best respond to their balance of reasons. If some courses of action are taken away by someone else

⁷ From now on, I am going to use the Razian term ‘exclusionary’ to refer to the kind of reasons practical authorities provide. However, I do believe that peremptory and exclusionary reasons are functionally the same: they both purport to exclude the subjects’ deliberation in some way.

(e.g., the authority giving exclusionary reasons), it seems as though agents' capacity to weigh those reasons effectively is negatively impacted.

If we think authority's exclusionary reasons negatively impact our practical reasoning, we might want to argue for a different view about authority. Those who contest the claim that authority provides exclusionary reasons often conceive of political authority as a theoretical authority. Theoretical authorities are experts on certain matters that can advise non-experts on what to do. Hurd (1991), for instance, argues that theoretical authorities' utterances provide reasons to believe that one ought to do what the authority advises. These are not, however, reasons to *do* what is advised. In this case, when a state tells its subjects to pay their share of taxes, for instance, the state is only giving individuals a reason to believe they ought to pay taxes. This reason for belief can prompt agents to act in accordance with the content of that belief, but it is not a reason to do as the authority commands. As a result, theoretical authorities' directives alter the subjects' balance of reasons by adding new ones. However, adding such reasons does not suspend the subjects' process of weighing reasons. Theoretical authorities' directives, which take the form of advice, only give information about the correct balance of reasons. They do not take the form of commands or orders. The evidence and advice the theoretical authority gives through its directives, then, affects individual practical reasoning differently than the practical authority's directives do. Authoritative directives exclude the subjects' deliberation and make them do what the authority commands. Advice from theoretical authorities gives evidence regarding what to do that affects practical reasoning only indirectly, for individuals can weigh the authority's advice and evidence against their first-order reason and consider that advice in deciding how to act. What Hurd is suggesting is that subjects are free to determine what their balance of reasons amounts to, via deliberation on the evidence they have. Authorities, through their expertise, merely shed some light on what the correct

balance of reasons is. Directives that only provide reasons for belief are, to Hurd, compatible with individual autonomy.⁸

Accounts that treat political authorities as instances of theoretical authorities do not work for two main reasons. Firstly, those accounts do not fully reflect how states operate and how their subjects think of them. States might give advice regarding which course of action to choose. For instance, they might, in conjunction with medical experts, advise citizens to get a seasonal flu vaccine to combat the disease's complications. In this case, the authoritative directive regarding the seasonal flu vaccine takes the form of expert advice. Individuals are free to decide for themselves whether to get vaccinated: they are not commanded to do so. However, as Shapiro (2004, p. 396) notes, political authorities mostly speak the language of obligation. They obligate their subjects to do as they order because they aim to make them act in certain ways. To do that, states provide their subjects with reasons for action, as opposed to reasons for belief. Reasons for action, that are treated as exclusionary, prompt one to act conclusively, whereas reasons for belief prompt one to consider those reasons for future actions. States want their subjects to act as they command; they do not want them to believe that they ought to act in a certain way, as that belief might not have enough force to make the subjects act on the directives.

Secondly, theoretical authority accounts cannot quite explain how states solicit a certain level of compliance with the law. Advice is meant to be considered and weighed against other factors and reasons. It might well be discarded if one decides to do otherwise. If authoritative directives were treated as mere pieces of advice, the subjects would have to deliberate whenever they are presented with a directive and decide whether to follow it. The consequence of this is that many might decide to discard the authoritative advice. This is an unpalatable

⁸ There can also be hybrid accounts that see authority as both theoretical and practical (Carr,1983).

element of those accounts that want to disregard SV. Reasons for action that exclude individual deliberation and prompt them to act as directed are better suited to explain how authorities solicit and ensure compliance (to a certain extent, at least), given that exclusionary reasons specifically direct the subject's conduct. Proponents of the theoretical authority model could of course say that states solicit compliance through threats and coercion. However, advice does not normally take the form of a threat. Advice is given to be considered, where individuals are free to make their own choices. It would be pointless for authorities to give advice backed by threats and sanctions when they could simply command.

SV, then, seems to be backed up by how political authorities operate, how they solicit compliance from their subjects and how their subjects think of them. If we think that the SV provides us with the most plausible explanation of the features and functioning of political authorities, we might want to attack Wolff's argument from a different direction. Another way of doing that is by rejecting P2's understanding of autonomy.

2.1.2 A Different Take on Autonomy

P2 argues that being autonomous (i.e., self-legislating), entails having a duty to give yourself moral commands. According to Wolff (1970), this duty is connected to the fact that agents take responsibility for their actions. Wolff claims that the duty to be autonomous lies in contrast with the duty to obey authority. When we do what the authority commands because the authority said so, regardless of the content of the authoritative directives, we do not reason and freely choose for ourselves the course(s) of action we want to undertake. In other words, when we obey authoritative directives, we are not giving ourselves moral laws. We are accepting as binding the laws someone else imposes on us. Hence, the contrast between autonomy and authority.

One might think that this conception of autonomy is too strong. ‘Autarkic’ autonomy, that is the Kantian conception of autonomy as “the self-sufficient formulation of one’s principle of action” (Bogg, 2002, p. 92) seems demanding and unsustainable. Wolff envisions the autonomous agent as someone who continuously strives to fulfil her duty to give herself moral commands. Following one’s moral commands might well be the opposite of abiding by someone else’s commands. But is autarkic autonomy sustainable in general? Agents cannot reach complete self-sufficiency in reasoning, deciding, and planning actions (Bogg, 2002). It would be extremely difficult and time-consuming, if not impossible, for agents to give themselves moral imperatives and plan their conduct at every turn. It could be more efficient and even desirable to delegate at least part of that planning and decision-making process to someone else.

It might be thought, then, that a less stringent definition of autonomy, one that does not see it as a moral duty of complete self-legislation, would be compatible with authority and, as a result, solve Wolff’s paradox. However, adopting a less demanding conception of autonomy does not make the tension between autonomy and authority go away. If we are not convinced by the strong, autarkic conception of autonomy as complete self-legislation, we might want to see autonomy as the exercise of self-governance, which amounts to responding to reasons and deciding for oneself how to act. This is a more nuanced sense of autonomy, according to which being autonomous simply means making decisions. You can incorporate other people’s practical reasoning into your own to make decisions, as opposed to having a moral duty to continuously give yourself moral commands.

However, treating authoritative directives as exclusionary appears to be in contrast with this modified notion of autonomy too. The exercise of the ability that rational agents have to decide for themselves and to act accordingly is hindered by the authority’s directives, even when we do not consider autonomy a moral duty. Responding to authority’s exclusionary

reasons seems incompatible with individuals deliberating and responding to reasons as a result of that deliberation. If we are autonomous, we can deliberate, whereas if we obey the authority, we cannot. We have seen that deliberating is important in terms of taking responsibility for our actions. This presents Wolff's dilemma again. Either we remain autonomous, or we obey the authority. The tension between autonomy and authority still stands. It is a tension between deciding for yourself by being responsive to your reasons and not being able to decide, due to being subjected to the authority and its exclusionary reasons.

2.1.3 The Autonomy-Authority Problem Reformulated

I argued that SV gives us the right definition of political authority. This means accepting that political authorities provide reasons that ought to be treated as exclusionary. If how we treat authoritative reasons collides with autonomy no matter how stripped down our characterization of autonomy is, then, autonomy and authority seem to be generally incompatible.⁹

Given that we specified the substance of P1 and offered an alternative view on autonomy, it is useful to reframe the argument for the autonomy-authority problem as follows:

P1*: Obeying the authority means treating authoritative reasons as exclusionary.

P2*: Autonomy entails deciding for yourself and responding to reasons that stem from that deliberation.

P3*: It is not possible to treat authoritative reasons as exclusionary while deciding for yourself and being responsive to reasons.

⁹ I discussed this argument on page 20.

C*: Therefore, authority and autonomy cannot coexist.

In the next few chapters (3, 4 and 5), I challenge P3* and show when and how the exclusionary element of authoritative directives and deciding for oneself are compatible. Prior to that, I turn to the analysis of three popular accounts of the legitimacy of political authority to see how they would respond to this reformulated version of the autonomy-authority problem.

2.2 Three theories of the legitimacy of political authority

This section analyses three influential accounts of the legitimacy of political authority. Section 2.2.1 deals with consent theories of legitimacy, section 2.2.2 is about membership-based accounts of legitimacy and 2.2.3 analyses an instrumentalist view on political legitimacy. Each sub-section starts with an overview of the theory in question, followed by some objections that have been raised against it or some problems the theory faces. I will then proceed to explain how the theory may approach the reformulated autonomy-authority problem and the solutions it may present to solve it. This is by no means an exhaustive list of prominent accounts of political legitimacy, and this analysis will inevitably leave out some important contributions made in this literature for reasons of space. I included three accounts that seemed to have engaged with the autonomy-authority problem, and that are closest to the kind of political authority I have in mind.

2.2.1 Consent theories of legitimacy

The first approach to analysing political legitimacy in this chapter is the one that grounds such legitimacy in the consent of the subjects. Classic consent theories argue that political authorities, such as states, are legitimate if and only if their subjects have validly consented to

them. Consent is valid when three conditions are met: a) the parties involved (consent-giver and consent-receiver) are competent; b) they know and understand what is going on and c) they consent voluntarily. Regarding a), being competent means possessing at least the basic mental capacities for practical reasoning, namely, the capacity to evaluate the pros and cons of different actions, and to form reasons for action and related intentions that prompt the agent to act according to her reasons. Infants, young children, people with severe cognitive disabilities and elderly people (in some cases) are usually regarded as incapable of consenting. Regarding b), when they have relevant mental abilities, agents can also know and understand what they are doing. Being competent and understanding one's situation and surroundings go hand in hand. Finally, one is thought to consent freely to something when c) that consent stems from one's free choices and actions. Acts of consent performed under duress, where the consent-giver is coerced into signalling her consent, are not instances of valid consent even when a) and b) are satisfied (Feinberg, 1989b).¹⁰

Before moving on to explain how consent works in legitimising political authority, it is necessary to say more about its definition and functions. Consent is thought to be an express or tacit act that enables two or more parties to modify the normative relationship between them (Kleinig, 2009).¹¹ These parties can be single agents as well as groups, such as organisations, institutions, and states.

Consent modifies the normative situation between the parties involved by altering the right-duty balance between them (Hurd, 1996; Kleinig, 1982). When one gives consent to

¹⁰ Thomas Hobbes (1651) is the only one, to my knowledge, who argues that consent given under duress constitutes a form of valid consent.

¹¹ An exhaustive analysis of all aspects of consent goes beyond the scope of this chapter. It is useful, though, to point out that there are two main theories of the ontology of consent, i.e., theories regarding what consent is. Some (Alexander, 2014; Hurd, 1996; Westen, 2004) argue that consent is only a mental state, ontologically independent of any act or signification that may accompany it. This is the mental state view on consent. Others (Dougherty, 2015; Feinberg, 1989a; Kleinig, 2009) state that consent is an act, something that has to be signalled from one party to another. This is known as the performative view on consent. Here, I treat consent as an act or performance.

another, one usually transfers certain rights, often acquiring obligations in return. These can be negative obligations not to withdraw consent without good reason, obligations not to interfere in the procedure consented to (e.g., surgery, or medical treatments) or positive obligations to obey the consent-receiver's commands if the consent-receiver is an authority one has given one's consent to.

Consent's function is, then, to be a deontic transformer. This function has two main 'applications', for the deontic transformation effected by consent can be used in interpersonal transactions and also play an exculpatory role in criminal liability (Feinberg, 1989a). As a deontic transformer, consent makes certain actions permissible that would be impermissible in its absence (Kleinig, 2009, p. 7). Firstly, to illustrate the case of interpersonal transactions, if A consents to B borrowing her laptop, A makes it permissible for B to perform an action that would not have been permissible without her consent. Had B just taken the laptop without asking, she would have stolen it and acted wrongly. In this way, consent changes the nature of the action itself by making it justified and morally and/or legally permissible. In our example, consent transforms what would have been a theft into a loan. These transactions operate against a background of rights and duties possessed by the parties involved. Consent helps to reshape the right-duty balance between the agents involved in the transaction. A has a property right that others do not use her possessions, while B has a correlative duty not to use A's property. By consenting, A alters the moral landscape and the moral relationship between her, and B. Consent allows A to temporarily waive her property right to enable someone else to use her property. Secondly, regarding its exculpatory role in criminal liability (Feinberg, 1989a), consent can make all the difference in cases in which a defendant (i.e., an alleged wrongdoer) is considered responsible or blameworthy for something she has done. If defendant D has the explicit consent of the complainant C to sexual intercourse, D will not be responsible for raping C, for C's consent, if valid, changed the nature of D's actions. Sexual intercourse is *prima facie*

wrong unless it is performed with the consent of the parties involved. When consent is not present, sexual intercourse becomes rape.

2.2.1.1 Political Consent

On the political side of the debate, I have said that when political authority is concerned, consent theories argue that there is only one necessary condition of legitimacy that has to be met by the state. The state has to obtain the valid consent of its subjects to be a legitimate authority. When consent is obtained, the political authority becomes a *de jure* legitimate authority. When those who are under the state's territorial jurisdiction consent to the state, the state acquires a right to rule over them that it can exercise through the creation of laws and directives. The subjects will have a duty to obey the authority they consented to.

Just as in the individual cases that are dealt with by private law (e.g., consent to sexual intercourse, to a medical examination and so forth), when citizens consent to the authority in the political case, consent changes the normative situation between the parties by altering the right-duty balance between them (Kleinig, 2009; Simmons, 2009). According to theories of this kind, before consent is given, the state has no right to impose obligations on its citizens, which means that the citizens are under no duty to obey its commands. Consent theories argue that the state acquires a right to rule and to impose obligations only after the citizens consent. This is not to suggest that individuals should not obey a political authority they have not consented to. There may be other moral or prudential reasons that justify obedience to a particular authority in certain (or, perhaps, most) cases. However, according to consent theorists, before the citizens' act of consent, the political authority is not legitimate, and cannot impose legitimate obligations on those who live in its territorial jurisdiction. The obligations that stem from the state's laws might be binding for those moral or prudential reasons, but are

not legitimate, if the consent of the governed is a necessary condition for the legitimacy of the political authority (Raz, 1995, p. 356).¹²

Generally, consent theories require the actual (as opposed to hypothetical)¹³ consent of everyone or most of those who are subjected to a certain authority (Simmons, 2009). This requirement has frequently attracted criticism. It is often objected that states cannot obtain the consent of most members of the political community. This would require unanimity, which is hard to reach in large-scale groups, such as political societies. Moreover, the high costs associated with soliciting and ascertaining people's consent, or with setting up mechanisms and procedures to solicit and ascertain consent where those mechanisms and procedures are not already in place, render consent an unattractive or infeasible way to legitimise political authority. In line with this objection, A. John Simmons (2009) argues that consent, where present, morally justifies the authority of the state, but that it is impossible to obtain in the kind of states we live in and therefore, no existent state is (or future state will ever be) morally legitimate.

To escape this objection, proponents of consent theories have often tried to modify the notion of consent to show that it is possible to obtain it from every member of the community. Some (Locke, 1689; Beran, 1987; Steinberger, 2004) argue that tacit consent, consent given by way of not opposing the authority, such as remaining silent or residing in a given territory, avoids the objection presented above.

However, not everyone in the political community always acts or omits acting in such a way that would signal consent. When this happens, the political community is virtually divided

¹² Locke is thought to defend this view. Nozick (1974) and Simmons (2001) are contemporary subscribers to the view that consent is a necessary condition for the legitimacy of political authority.

¹³ Hypothetical consent accounts aim to justify political authority by relying on what idealised agents in idealised conditions would agree to (Stark, 2000). Prominent hypothetical consent accounts of political legitimacy include Hobbes (2003); Rawls (1971); Gauthier (1986) and Scanlon (1998).

into two groups: those who signalled their tacit consent and those who did not. This leaves us with two options. We might think that those who did not signal their tacit consent are still expected to obey the state and treated as such, even though the state would not be legitimate for them. They might be punished if caught breaking the law, even though they did not tacitly consent to it, because their obedience is important to maintain public order. Or we could think that they should not be treated as having obligations to obey the state.¹⁴ On the one hand, if non-consenting citizens are still treated as if they had obligations to obey the law, there is the worry that tacit consent is useless. If, irrespectively of whether tacit consent is given, individuals are treated the same way (as if they had obligations to obey the law), we might think that this deprives consent of the normative force its proponents attribute to it (i.e., changing the right-duty balance between the parties involved and creating new rights and duties). If, on the other hand, we think individuals do not have obligations to the state when they do not signal tacit consent, and should not be treated as if they had, there is the question of how they should be treated. Can they be punished for their non-compliance with the law? Are they still full-fledged members of the political community? Or do they constitute a subgroup of citizens who, in virtue of non-consenting, is treated differently? These are important concerns that challenge approaches based on tacit consent. In chapter 7, where I discuss normative delegation, I argue that even though individuals are still obligated to obey the law when they do not delegate (or tacitly consent) to the authority, this does not deprive delegation (or tacit consent) of its normative force. I also discuss whether it is bad to act as though people have delegated to the authority and treat them as such despite the absence of delegation.

The most problematic aspect of tacit consent, though, is the difficulty of properly signalling it. Any act or omission can only count as tacitly consenting when everyone in the

¹⁴ Note that they might have moral obligations toward their fellow citizens or even moral obligations whose content happens to coincide with the content of the law.

community is aware that those actions or omissions will signify one's tacit consent to the authority. This is important because, no matter our view on the ontology of consent, consent has to be given intentionally to be valid. If we think that the act or omission constitutes consent, when the consent-giver does not understand the significance of her act or omission, she will not have consented. At the group level, if there are members in the political community who do not have the relevant understanding, these members' acts or omissions cannot count as consent. Voting is a relatively straightforward mechanism for ascertaining tacit consent than others. If the state makes it clear that voting is a way of seeking the citizens' consent, we can assume that most individuals understand what is asked of them when they vote, and it is easy to establish who did not vote. We would still have the problem of deciding how to treat non-consenting citizens, but we can set that aside for now. Silence and residency in a given territory are trickier in this respect. It would be more difficult to establish who, among those who are remaining silent or residing in a territory, has the right understanding of what their actions or omissions amount to. This makes ascertaining tacit consent difficult in large-scale groups such as political societies.

There is, however, an additional, more pressing, problem with tacit consent. Even if it were easier to ascertain tacit consent through silence or residency in a territory, and even if individuals knew that those acts or omissions signify consent, that consent might not be free. Hume (1748) argued that we cannot freely consent to a political authority because we are not presented with viable alternatives to consenting. The only options we have are tacitly consenting to the authority or emigrating to another country (and possibly subjecting ourselves to another authority). Emigrating, in most cases, is not feasible, due to the costs associated with it and the laws that regulate the procedure. Hume seems to think that consent is only valid when we have a variety of feasible options at our disposal. If our only alternative to (tacitly) consenting is emigrating, it seems that we do not have a choice but to submit to the state. This

could be generalised to say that whenever the alternative to consenting is sub-optimal, we do not have a choice but to consent. We can argue that we are still free to choose and not coerced into consenting even when our alternatives are suboptimal. Consider two scenarios. In the first scenario I am hanging over the edge of a cliff and I cannot manage to save myself. The stranger who is passing by can save me, but they demand that in exchange I give them my life savings. In the second scenario, a robber is holding me at gunpoint, demanding I give them my money or else they will kill me. What is the difference between the stranger asking for my life savings to save me, the robber holding me at gunpoint and the state soliciting my consent? If I do not consent to the stranger's request, I will be left to die; if I do not consent to the robber taking my money I will be killed and if I do not consent to the state, I am subjected to its laws (unless I emigrate). It seems that, in all these cases, there is no viable alternative to consenting.

However, there is a difference between these cases, and it amounts to who is responsible for one or more of our alternatives being worse than consenting. In the robber case, the robber is not only responsible for putting me in a difficult situation but also responsible for the fact that my alternative (i.e., being killed) is worse than consenting. The robber is constructing my choice set and forcing me to choose between consenting by handing out my money and being killed. This is a case of coerced consent, for, in a situation in which there is an imbalance of power (the robber has power over us exercised through violence), I am threatened into choosing something (handing out my money) where the consequences of non-compliance are so disproportionate that I have no choice but to submit. The stranger case is slightly different. Here, the stranger is not responsible for the situation I am in, for I slipped and, therefore, I find myself hanging over the edge of a cliff. However, the stranger is constructing my choice set in such a way that I feel compelled to choose to give them all my life savings over being left to die. We might argue that giving up my life savings is a better option than choosing to be left to die. Nevertheless, this might be a borderline case of coerced consent. The case of consent

solicited by a political authority might well be different. The state may not be responsible for the fact that exiting the state is much worse for us than remaining within its territory. And, unlike the robber and the stranger, it is not clear that the state is similarly constructing or altering our choice set. And even if the state were constructing our choice set so that the only alternative to consenting available to us is the hard one of exiting the state, as Leslie Green (1989) highlights, the choice to make might not have to be between consenting and leaving the state. It may well be that we can choose between consenting to the state, leaving or peacefully complying with the state's rules and cooperating without accepting its authority (Green, 1989, pp. 174-175). We might decide not to free ride, to follow the rules and enjoy the benefits of residing in that state's territory. This might be a case of insincere, tacit consent, where we go along as if we were consenting when, in reality, we do not take our action (or omission) as a way to signify consent.

Consent theories of political legitimacy have polarised theorists who either sing their praises or prepare to disprove them with a long list of objections. Consent theories incur difficulties with obtaining the consent of all or most group members. This prompted proponents to stretch the notion of consent to include more than actual, verbal acts as valid ways to signal it. But even the more inclusive notion of tacit consent seems to struggle to identify clear, unequivocal ways for most group members to signal their consent in political societies. The most attractive feature of consent theories, as we will discuss below, is their voluntarism, that is, the importance they attribute to the voluntary choice of everyone in agreeing to things. For this reason, these accounts seem suitable for preserving individual autonomy in political societies. I will argue that voluntarism does not always deliver the outcome we want (people obeying the authority) and that there are circumstances in which the consent or delegation of the citizens is not given, but authorities can act as though that consent or delegation has happened and treat their citizens as having consented or delegated. These are the cases

considered by urgent task theories. In Chapters 6 and 7, I will argue that coordinating people's conduct is a task of the utmost moral importance. For this reason, states can legitimately act as though their citizens have delegated their planning powers to them, even when citizens have not delegated, so long as their plans meet some conditions of adequacy (Chapter 6).

2.2.1.2 Consent and the reformulated autonomy-authority problem

It is time to turn to how consent theories may approach the reformulated autonomy-authority problem. We said that deciding for yourself (thus taking responsibility for your actions) in a way that is responsive to reason appears to be in sharp contrast with treating authoritative reasons and plans as exclusionary, which blocks and alters individual deliberation.

Consent theories would object to this point, thus challenging P3* of the reformulated paradox. P3* says that it is not possible to treat authoritative reasons as exclusionary while deciding for yourself and being responsive to reasons. Consent theorists could say that, by consenting, one acquires a moral obligation to obey the authority one consented to. So, obeying the authority one consented to does not violate individual autonomy because fulfilling one's moral obligation is a way to respond to reasons and, hence, compatible with autonomy. When one has an obligation to obey, doing so, might be argued, does not violate one's autonomy, as doing what one has a moral obligation to do is in line with responding to reasons. Having a moral obligation gives one a moral reason to act accordingly. Being autonomous entails, among other things, responding to reasons. Therefore, fulfilling a moral obligation means acting autonomously. When one is subjected to the authority one has consented to, one has a moral obligation to obey that can be fulfilled whilst remaining autonomous, as this obligation stems from a decision to voluntarily submit oneself to the authority.

This solution that consent theorists put forward to respond to the reformulated version of the paradox hinges on an element of voluntariness. It is the agent's choice to submit to the state that gives rise to the obligation to obey the authority. This voluntary choice is in line with individual autonomy as the exercise of the capacity to decide for oneself and is also a proper ground for political legitimacy, according to theories of this kind. It is worth noting that, for consent theories, without this voluntary choice to submit to the state, agents do not have obligations to obey the authority and whatever the authority imposes on them is a violation of their autonomy.

Considering agents' voluntary choices is important for their decision-making process, and certainly an attractive feature of consent theories. However, we can acquire obligations to obey the state even when we do not agree to it ruling over us. Moreover, following authoritative directives when we did not consent to the authority imposing them on us, does not always result in the loss of our autonomy.

Regarding the first point, we might sometimes find ourselves in situations where there are tasks of fundamental moral importance that need to be completed, and when there is someone or a group, that is well equipped to carry them out. Some argue that in these situations we have a natural moral duty to obey the commands of that authority that is in a much better position to achieve those tasks than us, whether we agree to it or not (Estlund, 2008).¹⁵ Think of those cases where a ship is about to sink, or a plane is about to crash. To save as many passengers as possible, a crew member starts imposing orders on any passengers to help out and do as she says. Do the passengers have a duty to obey the crew member's commands?

Consent theorists might think that we have a natural moral duty to rescue, where that duty gives us reasons to save people in cases of emergency if we can. But they might not think

¹⁵ A limitation of consent theories is that we do not straightforwardly decide what morality is.

that there is any duty for us to obey the commands of the crew member unless we agree to do as they say. For consent theories, the authority cannot impose valid obligations on us without our consent. But this voluntary dimension that consent theories deem necessary for others to legitimately impose duties on us, might not allow us to reach the goals we ought to have. If a ship is sinking, the priority is to try to save as many passengers as possible. This is what morality requires of the rest of the passengers and crew. But if passengers do not agree to do what the skilled crew members order them to do, the collective action required to save as many people as possible will not get off the ground, leading to the undesirable result that people will not be saved. In cases where there is an outcome of the utmost moral importance to reach, people's refusal to consent to the authority's instructions seems wrong. It seems that the urgency of a task (e.g., saving someone), coupled with agents stepping up and giving orders, can do some work in producing valid obligations for other agents to obey their orders. Soliciting and honouring individual consent, despite being so fundamental in several contexts, is not always necessary to give rise to obligations. If we think that we cannot impose obligations on others or have legitimate authority without the consent of the subjects, we risk failing to do the right thing in urgent situations that require people to impose coordination on others. In Chapter 7, I argue that achieving social coordination is a task of absolute moral importance and that states can legitimately impose obligations on their subjects to achieve it, even without their consent (when other conditions are also met).

But does imposing obligations on agents who have not consented to it violate their autonomy? The authority might have good reasons for demanding that other people around it do something, perhaps because the work it needs to carry out is morally important and requires a concerted effort to be executed. The authority might be legitimate in demanding something of others without their consent but, in doing so, it might temporarily damage their autonomous decision-making. We might think that, in such cases, agents might just have to do as they are

told and give up their autonomy for a while. However, this is not always the case. Agents do not lose their autonomy whenever someone else imposes obligations on them without their consent. Sometimes, obeying authoritative plans allows us to do a better job at acting on the correct balance of reasons than we would do if we spent time deliberating on what the correct balance of reasons is and how to act on it (Raz, 1986, p. 53). The authority might be an instrument for us to better respond to those reasons that apply to us anyway. In the ship case, we already have reasons to rescue as many people as possible when we are in a position to do so. By obeying a competent person or group that steps up to facilitate the rescue process, we respond to our reason to help people in need and fulfil our obligation to rescue others better than we probably would have if we had tried to manage this big task by ourselves. In doing so, we do not lose our autonomy. Instead of planning and executing the rescue ourselves, we rely on someone else to do the planning and execute the rescue anyway, albeit differently.

This is how we can respond to the worry that discharging duties imposed on us without our consent constitutes a violation of our autonomy. It is worth noting that Wolff (1970) would argue that voluntary acts of consent constitute a forfeiture of individual autonomy because it is not admissible, in his eyes, to treat authoritative reasons as exclusionary. I already defended accounts of practical authorities that operate by providing exclusionary reasons. I will analyse exclusionary reasons further when I discuss Raz's account in section 2.2.3, after first exploring Wolff's worry further.

Wolff is worried that acts of consent themselves might violate autonomy. Consenting to obey another means subjecting oneself to the authority of another that makes decisions on our behalf. While delegating part of our decision-making process to someone else is a way to respond to reasons and does not necessarily impair autonomy, as we said before, there are risks associated with it. The most obvious one is when the person we relied on for planning on our behalf goes rogue and acts immorally or in ways we did not authorise. We have, then, the

problem of identifying when the authority is acting inappropriately. We need a mechanism in place that allows us to check whether or not the authority is acting appropriately, where this can be public reasoning on the part of the authority itself to disclose its operations. So, in this respect, we can charitably interpret Wolff's worry as the worry that consenting might not suffice as a ground of legitimate authority. Perhaps we need some way to check on what the authority is doing that allows us to recognise when the authority's actions impair our autonomy.

At this point, we might think the problem of the authority violating our autonomy or breaching the terms of our mutual agreement by acting immorally can be solved by withdrawing our consent. However, withdrawal is not a straightforward process in consent theories.

When we consent to the state at a time t , we agree that the state will make decisions on our behalf in the future. We establish we are giving our consent, and after consenting, we do not get to decide on matters regulated by the state anymore. We agree to follow the laws the state imposes on us and, if we think the law is wrong or that we should no longer obey the state, we have to withdraw our consent. However, withdrawing consent is not easy. There are no clear procedures as to what is involved in actively dissenting and how the state should behave towards dissenters. The thought here is that an act of consent given at t subjects you to a lifetime of the state making decisions on your behalf that morally obligate you. If you could freely cancel your consent and obligation to obey whenever you think rationality requires you to do so, this would not be a problem for individual autonomy. A simple decision could reverse your act of consent. However, consent is never cancellable in such a strong and neat way. If you could only cancel your consent regarding future commands, for instance, and not commands that have already been issued, by consenting you would partially lose your autonomy. You would not have the possibility to immediately reverse your initial decision (consent) and this massively impairs your ability to decide for yourself what to do.

It is worth noting that not being able to reverse your decisions can be irrational in some cases. Rationality requires that individuals act on the balance of reasons and, on some occasions, it requires them to quickly change their mind regarding what to do due to new circumstances or new reasons that have arisen. Consent that cannot be immediately cancelled when necessary, may drive individuals that have changed their minds to act against the balance of reasons. If they cannot immediately withdraw their consent, they still have an obligation to obey current laws. If they think a law goes against reason, they might decide to obey it anyway because the other alternatives available are too costly for them. If they do obey that law despite changing their mind, they would act irrationally, for rationality requires us to respond to reasons.

To conclude this section, I believe consent theories have one main advantage and two disadvantages. They emphasise voluntariness, where the agent's choice to submit to the state is considered a proper ground for political legitimacy. This appears to be in line with individual autonomy as the exercise of the capacity to decide for oneself. If the balance of reasons requires it, agents can decide to delegate certain tasks to the state and, in doing so, they exercise their autonomy. However, despite the importance of voluntarily choosing to obey authority, voluntariness is not the only valid ground for obligations. The authority can legitimately impose plans and decisions on us, for instance, when there is something of utmost moral importance to do and our cooperation is required. Consent theories seem to neglect this ground of obligations when they argue that consent is the only necessary and sufficient condition of legitimate authority. Moreover, the difficulties associated with withdrawing one's consent and their implication on individual practical reasoning, make consent too rigid. Acts of consent that are hard to reverse or cancel do not allow the subjects to freely exercise their practical reasoning.

2.2.2 Membership-based accounts of legitimacy

We previously discussed consent theories, which are (for the most part) voluntarist theories of legitimacy and obligation. According to these theories, it is the voluntary choice¹⁶ of individuals that grounds things such as obligations or the legitimacy of political authorities. Membership-based accounts are different in this respect. They can be classified as non-voluntarist theories of legitimacy. Those who appeal to membership as a source of obligations believe that there can be duties imposed on someone regardless of that person's free choice or voluntary undertaking.

Groups or communities have their own rules, regulations, laws or norms. Being a group member entails abiding by the rules, regulations, laws or norms of the group and doing one's part toward the realization of the group's enterprise(s). Therefore, group members have to comply with the group rules and regulations and do their part toward the common enterprise(s). To put it differently, duties derive from and are attached to group membership (Dagger & Lefkowitz, 2014).

In some cases, groups can be joined by their members voluntarily, that is, individuals can freely decide whether or not to become part of a group, such as in the case of joining a sports club or becoming a volunteer for Greenpeace and so forth. In other cases, individuals just happen to be members of some groups, without them having done anything to signal their desire or willingness to join those groups. This is the case, for instance, of families, where family members do not, at least in paradigmatic cases, do anything to join the group in question and are simply born into a particular family. When one freely decides to join Greenpeace as a volunteer, that is sufficient for one to acquire some obligations that stem from being a member of Greenpeace, such as a duty to carry out some environmental projects for the company and

¹⁶ Except for hypothetical consent.

the duty to attend volunteer meetings. Similarly, the fact that a person was born into a family is sufficient for that person to acquire familial obligations (i.e., those obligations that arise out of being part of a family), such as abiding by her parents' commands as a child or taking care of her sick parents as an adult. Parents also acquire duties to provide for their children when (and most likely even before) they bring them into the world, where these duties derive from the place that the parents occupy in their family.

Regarding political authority, just as in the case of familial obligations, obligations are legitimately imposed on citizens due to them being members of a specific political society or community. Individuals, in most cases, do not join states voluntarily. They happen to be born in a political society and their being part of that society is sufficient, proponents of the theory say, to ground legitimate obligations. In this case, the legitimacy of a state stems from the fact that those who are subjected to it must obey its laws *qua* members of the political community regulated by that state.

Membership-based accounts of legitimacy can take many forms. An influential theory based on membership is the one advanced by Ronald Dworkin (1986).¹⁷ Dworkin claims that membership in or association with a political society grounds political obligations. He defines these obligations as associative or communal obligations (Dworkin, 1986, pp. 195-197). These have a lot in common with those obligations that are attached to the roles that people play in families and friendships. According to Dworkin, both kinds of obligations, political obligations and familial and friendship-based obligations, cannot arise from a prior agreement or consent. Moreover, associative duties, much like familial and friendship-based duties, are related to social roles that people play in their life. Dworkin believes that four conditions have to be met for associative or communal obligations to be valid in political societies. (1) Group members

¹⁷ For more membership-based accounts see Gilbert (2006); Horton (2006) and Scheffler (2018).

need to consider themselves as having special obligations (or responsibilities) to other group members; (2) associative obligations are owed to each of the other group members; (3) associative obligations are based on a concern for the other group members' wellbeing; and, finally, (4) that concern has to be equal for every member of the community (Dworkin, 1986, pp. 199-200). Political societies or associations that satisfy these conditions legitimately obligate their members to abide by the rules or laws of the group. Dworkin calls communities that satisfy the four conditions 'communities of principle', where members see themselves as bound by principles they share with the other members.

We might think that the analogy with familial and friendship-like bonds is inadequate to capture the functioning of the legal model of obligations. We might think that because, like Simmons (1996), we believe that the strong bonds and affection that we might develop for friends and family members are too dissimilar from the feeling (or utter indifference) that, *qua* citizens, we may feel toward the state or our fellow citizens. We can develop some form of psychological bond with other citizens, a citizenly tie that stems from being part of the same political community. But that seems very different from familial and friendship-like bonds. Dworkin's position can be defended by thinking of the analogy with familial and friendship bonds differently. What Dworkin might be arguing is that equal concern for all our fellow citizens is a moral standpoint that should guide our (and policymakers') thoughts and actions. An equal concern is what citizens are due as members of their political community and, in thinking about what to do, we should settle on actions that do not betray or violate that concern. We think that all our family members and our friends deserve our concern and, arguably, to be treated equally. Similarly, our fellow citizens deserve our concern, and we should be impartial and equally concerned about them all when it comes to matters of public affairs.¹⁸ This does

¹⁸ This is similar to what relational egalitarians such as Anderson (2010); Kolodny (2014) and Viehoff (2014) argue for.

not mean that the concern we have for other members of the political community has to match how we feel about family and friends in the intensity and nature of the bond. It simply means that the psychological bond we form with members of our community has to be felt equally and impartially by everyone.

What if one does not feel equal concern for one's fellow citizens? One might be indifferent to them and have no feelings of any kind towards them. This may happen because not everyone will feel a sense of belonging to the political community. We might think that Dworkin's theory has the consequence that those who do not feel equal concern for fellow citizens are not bound by the laws of their community. Dworkin (1986) responded to similar objections by arguing that citizens do not have to feel or be aware of that special relationship with other group members. What matters is that their behaviour can be interpreted or seen in a way that satisfies the aforementioned conditions. For instance, the behaviour of someone who abides by the laws of her country and does not do anything that violates other group members' rights, could be seen as the kind of behaviour that stems from a special relationship or bond and equal concern that that person has (without realising) for her fellow citizens. Dworkin (1986) remarks that citizens must see political obligations as stemming from a general responsibility everyone has of concern for others' well-being. Political obligations are attached to the fact that people ought to have an equal concern, not to them having that concern.

Following the overview of membership-based accounts of legitimacy, it is now time to inquire how these theories would solve the reformulated autonomy-authority paradox I presented previously.

Associative obligations, which arise out of membership to some group, constitute an interesting puzzle to Wolff's (1970) account of the relationship between autonomy and authority. There is a threat to Wolff's argument posed by the possibility of obligations that are

not entered into voluntarily. Wolff seems to admit that if I consent to a political authority, I, then, have an obligation to obey the authority. He, then, acknowledges voluntary obligations, which we acquire when we consent to certain things. His main point, which constitutes the basis of his philosophical anarchism, is that we should not consent to obey political authorities because doing so violates our autonomy. For Wolff, we have the power to consent to the authority governing us but, in consenting, we forfeit our autonomy and accept that someone else will now rule over us and make decisions for us, which is an undesirable result. His anarchist argument concludes that we should not voluntarily enter into an obligation to obey¹⁹ and that, even when the actions of others can give us reasons and alter our reasoning in different ways, they do not give rise to exclusionary reasons for us to follow.

But what if we had obligations to obey that do not depend on my voluntarily entering them? If that were the case, and we saw that there can indeed be non-voluntary obligations, then the fact that I should not voluntarily enter such obligations is irrelevant. Wolff seems to face the following choice: he could either (1) deny that any obligations stem from involuntary group membership, or (2) admit that there are such obligations but argue that an obligation to obey commands (from the authority) cannot be one of them.

Regarding (1), it seems difficult to deny that obligations arising from group membership exist. It seems intuitive that we can acquire obligations by virtue of occupying a role with certain functions within a group, whether we agreed to take on that role or not. We can think of roles themselves as being constituted by the obligations that role holders owe to other agents that are connected to them (Dagger & Lefkowitz, 2021). Being a parent means that you owe obligations to your children, obligations that you acquired the moment you became a parent, whether that stemmed from a voluntary choice or not (Archard, 2010; Prusak, 2013).

¹⁹ Presumably there will be exceptions to this stringent requirement that Wolff poses but they do not seem to be explored much by him.

Conversely, being a role holder might come with significant benefits that are received by virtue of being part of a group. Children might reap some benefits *qua* family members, as their parents provide them with the financial, emotional and physical stability they need throughout their formative years. Taking advantage of the benefits that stem from being a member of a group can also give rise to obligations towards the other group members. In this case, children might acquire obligations to care for their elderly parents down the line (Keller, 2013; Lee, 2015; Stuifbergen and Van Delden, 2011). It would be difficult for Wolff to deny that we can have associative obligations as group members in similar cases.

Wolff, however, could try to argue for (2) and say that obligations to obey the commands of the authority are not associative obligations. We might acquire obligations to obey the authority or obligations to behave a certain way towards our fellow citizens when we voluntarily choose to do so. However, unless we agree to it, which is something we should not do anyway for Wolff, we do not have associative obligations to obey the authority. How to respond to this objection? Dworkin (1986) insisted on the analogy between familial obligations and state obligations, arguing that associative duties are related to the social roles that people play in their life. Being a parent, a child or a citizen of a political society are all roles to which certain responsibilities, claims and duties are attached. If this analogy between the family and the state holds, then we can have obligations that arise from non-voluntary membership of a state. Citizens can have duties to obey the law of the state they live in because they were born into this role which, as Dworkin argues, requires that their actions exhibit an equal concern for their fellow citizens. Treating everyone with equal respect is just what being a citizen requires, whether we chose to be citizens of that particular country or not. Citizens might also benefit greatly from being part of a political community that takes care of them and provides them with valuable services. And this might be a way for them to acquire duties towards their states or fellow citizens. One might not find the family-state analogy convincing. But if roles, whether

they are familial, social, political or professional, function similarly in different domains, it is plausible that duties are attached to them whether they are undertaken voluntarily or not.

Why would Wolff be so suspicious of non-voluntary associative duties? Non-voluntary associative duties might be problematic when membership in a group is a sufficient condition of legitimacy. Some states exercise their political power morally while some others occasionally or consistently exercise them immorally. A political society governed by an evil tyrant, for instance, is an example of a state that is run immorally by its rulers. The tyrant exploits its citizens, passes laws that violate human rights and even imprisons a sub-group of citizens based on their ethnicity. If membership in a group is sufficient to ground legitimate obligations, even members of badly administered groups such as the one just described might be bound to obey the group's rules, despite the content of the rules being immoral. In this evil tyrant society, it is hard to think that individuals are autonomous. The subgroup of citizens who are discriminated against and imprisoned are deprived of the possibility of deciding for themselves and cannot be responsive to reasons. And if the laws go against the subgroup of free citizens' balance of reasons, these citizens are not autonomous when their group membership bounds them to obey those immoral laws.

The evil tyrant state, though, violates Dworkin's conditions of the state and citizens treating each other as equals. When membership is not a sufficient condition of legitimacy, appropriate associative obligations do not arise for group members if the state does not meet certain criteria. Membership in a group might hinder individual autonomy in some circumstances. However, when membership is not sufficient to ground legitimate obligations, we need some other conditions that, coupled with membership, produce legitimacy. Dworkin's theory stresses the importance of the non-voluntary element of political obligations but most of the normative work when it comes to legitimacy and legitimate obligations is done by the principles adopted by the political community people are members of. In Dworkin's case, it is

not just membership that grounds obligations. It is membership coupled with the principles of equal concern that a society adopts. The implications of his theory, then, do not directly result in a violation of individual autonomy.

We might disagree on what conditions, aside from membership, we need to ground state legitimacy or legitimate obligations. We might think, like Dworkin, that political societies need to treat their members as equals, or we might require, aside from membership in the community, that the laws of the community follow the principles of justice or a combination of these requirements. If we look past a potential disagreement over what other conditions are required, we will see that what matters is that the state produces plans for its subjects that are adequate by certain standards we set. These standards might deal with treating people as equals or abiding by the principles of justice. What is important is that whatever standards are set by the political community, the state puts forward plans that conform to those standards. In Chapters 6 and 7, I argue that membership matters when it comes to deciding which authority to obey. If one is born in the UK, one acquires the role of a UK citizen at birth, and this is important to determine which authority one needs to respond to. The role of UK citizen requires that one abides by the laws of the UK, not the US unless one has ties in both countries. But we need something more than membership for state legitimacy. I argue that we need the state to put forward adequate plans that are sensitive to solving societal problems and give public justification for them.

2.2.3 The Instrumentalist View

The last group of theories of the legitimacy of political authority I consider here is known as the instrumentalist view of legitimacy. The most influential account that falls into this category is Joseph Raz's (1986, 2006) service conception of authority. In this chapter, I will focus

exclusively on the service conception, leaving out other accounts that classify as instrumentalist accounts of legitimacy.²⁰ I start with an overview of Raz's theory and finish with an analysis of a Razian response to the reformulated autonomy-authority paradox.

According to Raz (1986), an authority is legitimate when it serves its subjects, namely, when it helps the subjects to better conform to the balance of reasons. Raz believes that political authority should facilitate the process of practical reasoning for its subjects. The authority does so when it adheres to three conditions, which form the bulk of the service conception: the *dependence thesis*, the *normal justification thesis* (hereafter NJT) and the *pre-emption thesis*. The dependence thesis (Raz, 1986, pp. 40-47) states that authoritative directives are legitimate when they are based upon reasons the subjects have independently of the authority and its commands (dependent reasons). This means that a political authority specifically tailors its laws to reflect those dependent reasons that the subjects have independently of the law. The NJT argues that, for a directive to be legitimate, the subjects need to do better at conforming to the balance of dependent reasons by following authoritative directives than by trying to decide for themselves what the balance of dependent reasons requires (Raz, 1986, pp. 53-54). In other words, the authority, through its directives, should help the subjects to better conform to first-order reasons. Finally, the pre-emption thesis stems from the dependence thesis and the NJT. This is a thesis about how subjects should treat authoritative directives (Ehrenberg, 2011, p. 887). For Raz, the authority's reasons pre-empt the subject's reasons for action in the sense that, when the authority requires the performance of an action, that request is a reason for doing that action that excludes and replaces other reasons the subject might have for not so acting (Raz, 1986, p. 46). The subjects should, then, allow authoritative directives to be rules that

²⁰ For other instrumentalist accounts of legitimacy see R. Arneson (1993, 2003) and S. Wall (2006). These are instrumentalist accounts of *democratic* legitimacy, which focus on democratic political authorities.

guide their actions, and this requires treating those rules as motivating reasons (Ehrenberg, 2011, p. 887).

The reformulated version of Wolff's paradox that I presented in section 2.1.3 sees autonomy and authority as incompatible because how we treat authoritative directives (as providing exclusionary reasons) seems to collide with deciding for oneself autonomously. How would Raz solve this reformulated autonomy-authority paradox? For Raz, autonomy and authority are not irreconcilable. There are instances in which individuals can be autonomous despite being subjected to political authority. This happens when the authoritative directives are based on the subjects' dependent reasons. By complying with directives that satisfy the dependence thesis, the subjects would respond to reasons that already apply to them, thus complying with their balance of reasons. Responding to reasons is partially what makes one autonomous, so the service conception at least partially solves the problem of how individuals can be autonomous in the face of authority.

Furthermore, for Raz, the duty to obey the authority can be imposed on agents regardless of their choice. The authority does not need the subjects' consent to impose laws on them. This gives Raz's theory an advantage over consent-based accounts of legitimacy. For consent theories, we are autonomous despite treating authoritative directives as providing exclusionary reasons only when we consent to the authorities imposing those directives on us. Raz's account avoids the difficulty associated with obtaining the consent of the majority in political societies. Whenever the authority's directives are based on the subjects' dependent reasons, the subjects maintain their autonomy in abiding by them whether they consented or not.

Raz's theory accounts for the first component of autonomy, namely, responding to reasons. It does so, though, only when authoritative directives are based on pre-existing reasons the subjects already have. Directives that are not based on the subjects' pre-existing reasons

are not legitimate in the service conception. However, it might be the case that the authority bases its directives on reasons that do not already apply to the subjects. Are those directives illegitimate? Are the subjects allowed to simply discard them? Authorities might help us comply with our balance of reasons whether their plans are based on our pre-existing reasons or not. We can take into consideration the authority's directives in planning our own conduct and use them as a way to fulfil our goal and do what we have the most reason to do. What matters is that we respond to our balance of reasons when we abide by the law. If we can do so even when authoritative directives are not based on our dependent reasons, we would be able to exercise our autonomy in a variety of different contexts. In Chapters 5 and 6 I argue that we can incorporate the authority's reasons into our decision-making processes and practical considerations regarding what to do. When we follow authoritative directives and plans, we are able to respond to reasons more efficiently. Moreover, authoritative planning can expand our reason-responsiveness capacity because authoritative reasons can guide our conduct the same way our reasons would.

So, we can follow Raz in thinking that it is a good desideratum of an account of legitimate authority that the authority should allow us to respond to reasons through their directives, as responding to reasons is a prominent element of autonomy. But we can go a bit further in arguing that this can be the case even when authoritative directives are not directly tracking the subjects' reasons. This accounts for the first component of personal autonomy more fully.

What about the second component? Can the subjects decide for themselves under the service conception? Under the service conception, agents can decide for themselves how to act up to a certain extent. Treating laws and directives as exclusionary can make some room for individual deliberation. Political authorities, through their laws and directives, provide parameters of action for their subjects, especially in the context of large-scale groups. Social coordination is necessary for the functioning of political societies and the fulfilment of personal

goals. Without that kind of coordination and social order, individuals would not be able to act with others and complete their projects. Authoritative directives, when followed, provide the normative and legal background against which we can act. Deciding for oneself does not amount to creating every plan that pertains to us or making decisions about every single aspect of our personal and social life. It amounts to being able to make choices and decisions within the parameters required by the social contexts and groups we are part of. Treating laws and directives as exclusionary can make room for individual deliberation.

Raz's account, though, is subject to the same criticism as consent-based theories of legitimacy. I said that consent-based theories might lack a proper 'escape clause' that would allow the subjects to withdraw their consent to the authority and do otherwise. Similarly, the way exclusionary reasons work within the service conception does not allow for the subjects to re-open deliberation freely if needed (e.g., when the authoritative directives do not enable them to respond to reasons or are immoral). Since exclusionary reasons heavily affect individual deliberation, we need a reliable mechanism for agents to understand when they can disregard a command. There are instances in which our first-order reasons oppose the reason(s) provided by a directive. How do we know what to do in those instances? To decide what to do in cases in which our first-order reasons contradict the reason(s) provided by the authority, we would need to deliberate and weigh our reasons against authoritative reasons. However, when authoritative reasons are taken to be second-order, exclusionary reasons, those reasons prevent us from choosing a course of action that goes against the content of the directive. For instance, a soldier who has been ordered to do something that goes against his balance of first-order reasons needs to take that order as a reason not to act on his view on the matter (Perry, 1989). Exclusionary reasons enter our practical reasoning and block (at least partially) our deliberation by giving us reasons to refrain from acting for some first-order reasons. They block courses of

action that would go against the directive so that we choose to follow it rather than disregard it. How do the agents know when to act otherwise, then?

Raz's solution is that we should determine in advance clear cases of 'faulty' exclusionary reasons, namely, cases in which we ought to ignore authoritative directives. This way, we will know the kind of directives that fail to provide exclusionary reasons and will, therefore, do otherwise. However, it would be difficult, if not impossible, to determine in advance all possible scenarios in which we are required to break the law. Likely, the subjects could often be in the presence of faulty directives and reasons without knowing it. Given that their ability to decide and to do otherwise is blocked by authority's exclusionary reasons, the subjects could find themselves complying with laws they should not follow, without being able to properly exercise their practical reasoning. Moreover, the authority is the only one that can decide when it is permissible to break the rules. There is virtually no room for agents to autonomously decide when to go against the law. For agents to fully exercise the decision-making capacities that form part of their autonomy, they need to be able to re-open deliberation if they have reasons not to act on the authority's reasons. Chapter 6 argues that a better 'escape clause' to bad exclusionary reasons is public reasoning on the part of the authority. This allows the subjects to freely check on the authority's operations if needed and make informed decisions on whether to re-open deliberation on authoritative directives.

Re-opening deliberation when stringent reasons to do otherwise arise is an important aspect of practical reasoning. As Michael Bratman (1999a) shows us, being able to change your mind if needed is just as important as settling the matter on what to do if we want to act rationally. Not being able to change one's mind can be irrational. When an agent plans their conduct, the plans they makes are pro tanto, they can be modified if necessary. This enables the agent to quickly adapt to their environment and to maximize their exercise of practical reasoning. Agents should be able to disregard the law if new circumstances rightly require them

to do so. Establishing in advance all the cases in which exclusionary reasons can be ignored would be difficult. To effectively exercise their practical reasoning, agents should be able to decide for themselves whether to disregard some exclusionary reasons.

As we will see in Chapter 3, where I analyse Bratman's account of individual practical reasoning, in individual cases of planning agents can re-open deliberation when needed. Similarly, in cases of small-scale shared agency, agents can discuss with others whether deliberation should be re-opened on some matters. In cases where directives come from above, it is more difficult to establish when it is appropriate to reconsider and do otherwise. Aside from established cases of faulty exclusionary reasons that blatantly fail in pre-empting individual conduct, in Raz's theory, agents cannot *decide to go* against authoritative directives of their own accord. The service conception only enables them *to go* against authoritative directives when the law says so. To fully exercise their capacity to decide for themselves, agents need to be able to discern when to reconsider. I will explore how reconsideration works in individual, shared and hierarchical agency in the next chapters.

2.3 Summary

Raz's account of legitimate authority gets very close to giving us a full, satisfactory picture of how autonomy and authority can coexist. His theory accommodates the first component of autonomy, namely, responding to reasons. But it does so only when authoritative directives are based on pre-existing reasons the subjects already have. Furthermore, agents can decide for themselves under the service conception but only to a certain extent. This account does not allow for the subjects to re-open deliberation freely if needed because only the authority that can determine faulty cases of exclusionary reasons that the subjects can discard.

In determining when political authorities are legitimate, I will closely follow Raz's account as it seems to be the one that approximates what the right balance between authority and autonomy should be (compared to the other two accounts I previously analysed). In the chapters that follow, I explore the possibility that directives that are not based on the subjects' pre-existing reasons are legitimate and help the subjects to act on their balance of reasons. This would allow individuals to exercise their reason-responsiveness capacities regardless of whether the authority takes their pre-existing reasons into consideration in creating laws. Moreover, I argue that public reasoning on the part of the authority is needed because it allows individuals to check on the authority's operations when second-order reasons arise to go against authoritative directives. This mechanism allows individuals to know when to re-open deliberation on certain matters and do otherwise, thus not taking authoritative reasons as action-guiding anymore.

The fact that authoritative reasons need to be taken as action-guiding is an important aspect of Raz's account. When we take a personal plan or someone else's directive as action-guiding, that directive or plan features prominently in our practical reasoning. It is easy to treat our reasons as action-guiding or motivating. We represent those reasons in our minds, and they directly interact with our other reasons for action, our desires, beliefs and so forth. But how can reasons that belong to someone else come to guide our conduct in the way our own reasons would? Can we truly take the reasons that stem from authoritative directives as a guide for our actions in a way that goes beyond mere obedience to those directives? In Chapters 5 and 6 I take seriously Raz's idea that the reasons that stem from authoritative directives can become our motivating reasons.

Chapter 3 - Individual Practical Reasoning, Plans and Autonomy

3.0 Introduction

In Chapter 2, I examined Wolff's (1970) autonomy-authority problem, also known as the paradox of autonomy and authority. I also analysed three popular accounts of the legitimacy of political authority.

Wolff's version of the paradox does not illuminate the complex dynamic between autonomous agents and the authorities they are subjected to. His account seems to interpret the tension between autonomy and authority as the tension between running your own life in a certain way and the authority running your life for you through commands and directives. But why is it preferable to be in charge of your own life rather than letting someone else take over it? And why can we not just find a balance between obeying authority and deciding for ourselves how to act?

However, a detailed analysis of Wolff's argument reveals that the tension between autonomy and authority is ultimately the one between deciding for yourself (thus taking responsibility for your actions) and not being able to decide for yourself because the authority (partially) blocks your practical reasoning. Deciding for yourself is connected to moral responsibility and, for that reason, it is fundamental to investigate whether individuals can still be autonomous and take responsibility for their actions when subjected to authority.

Ultimately, what seems to cause a clash between autonomous decision-making processes and following authority is the fact that we treat directives as exclusionary reasons. Treating authority's directive (and reasons) for action as exclusionary and deciding for yourself in a way that is responsive to reasons are *prima facie* incompatible.

In the second part of Chapter 2, I investigated how three influential accounts of political legitimacy would solve the reformulated version of the paradox I presented and argued that Raz's (1986) solution comes closest to solving it. Raz's theory explains how individuals respond to reasons under the authority. But this is only possible when authoritative directives are based on pre-existing reasons the subjects already have. Furthermore, in his theory, agents cannot re-open deliberation freely when needed because it is the authority that determines when it is permissible to do so. These limitations warrant further investigation to see whether agents are still autonomous when authorities do not base their laws or directives on their reasons. Moreover, re-opening deliberation when needed seems essential to practical reasoning, as being able to change our minds if needed allows us to act rationally. In this chapter, I highlight the importance of re-opening deliberation in cases of individual practical reasoning.

So far, I only mentioned a *prima facie* incompatibility between autonomy and authority, as the literature on the matter primarily sees a tension between agents deciding for themselves and the authority making decisions for them. There is, however, a *prima facie* general tension between autonomy and shared agency that stems from the fact that any kind of plan, not just plans made by authorities, has an exclusionary nature.

By shared agency I mean those instances when individuals do things together (with or without obeying an authority). Agents often act with others. In contexts of shared agency, individuals create a shared plan (Bratman, 2009) toward common goals. When they do so, they have to consider other people's choices and reasons for action. Moreover, shared plans need to be treated as exclusionary for agents to be able to stick to them. In these egalitarian settings with no authority, one might have to respond to reasons that are not the product of one's practical reasoning. They might not even be reasons one had before adopting the shared plan. It needs to be clarified then whether agents are deprived of their autonomy or prevented from fully exercising it both in egalitarian (e.g., performing joint actions with equals) and

hierarchical contexts (e.g., the legal case of obeying the state). The tension between autonomy and authority then seems to be part of a more general tension between autonomy and shared agency.

I want to argue that individuals can remain autonomous in egalitarian and hierarchical cases of shared agency despite abiding by other people's reasons and treating shared plans as exclusionary. To show how agents remain autonomous in these contexts, I propose a three-stage model of practical reasoning which analyses all relevant moments of decision-making, from acting individually to acting with others and complying with authority. The model highlights the exclusionary nature of plans and how our autonomy can be preserved despite treating plans as exclusionary in different contexts.

In this chapter, I begin the analysis of a three-stage model of practical reasoning which will solve the autonomy - shared agency problem. Here I focus on the first stage of the account, that is, individual practical reasoning. This chapter examines the mental processes behind individual deliberation and decision-making. It explains how agents' various mental states combine to direct their conduct. These mental states, with emphasis on individual intentions and plans, in conjunction with norms of practical reasoning, enable agents to decide for themselves what to do and to be responsive to reason. In other words, they enable agents to act autonomously. The chapter consists of four main sections after this introduction. Section 3.1 gives an overview of the main mental states involved in individual deliberative processes, with a focus on plans. Section 3.2 provides some explanatory models of how plans and the related psychological processes lead us to action and a general account of individual practical reasoning. Section 3.3 concludes the analysis of Stage 1 of the three-stage model to solve the autonomy – shared agency – problem. It investigates whether the norm of practical rationality according to which we must treat plans as exclusionary threatens individual autonomous agency. Section 3.4 summarises this chapter.

3.1 Intentions and Plans

In this section, I define the main mental states involved in agents' deliberative processes. The focus will be primarily on goals, intentions and plans. These mental states combine with beliefs and desires in the deliberative processes of a rational agent, motivating and causing her to act in specific ways.

Beliefs are attitudes we have when we take something to be the case or to be true (Schwitzgebel, 2021). When I believe that it is raining outside, I take it to be the case that it is raining outside, because I see the raindrops forming puddles outside my window or because someone tells me that it is raining. When it comes to desires, it is thought that we desire something when we are disposed to act in such ways that would bring about what we want (Schroeder, 2020). If I desire to go sing in the rain now, perhaps because I believe it is raining now, I am disposed to take the relevant steps that would bring about this desired state of affairs of singing in the rain now. We can think of goals as the contents of our final desires, or perhaps the properties of the world we value that our desires are responding to. Perhaps I value feeling the fresh raindrops on my face (goal) and this explains why I am disposed to act in relevant ways (i.e., why I desire to sing in the rain). Beliefs and desires can produce intentions and plans, which allow us to act in relevant ways. Even though I am disposed to act in a way that brings about singing in the rain, I do not always act in such a way that would bring that about. Maybe I am in a work meeting, and it would be frowned upon if I would just leave to go sing in the rain. Some think that sometimes we need an additional commitment to action that contributes to producing it. This commitment to action is an intention.

If I intend to walk my dog, I am committed to doing so, where this commitment prompts me to organise my conduct in a way that I bring about that I walk my dog. Intentions are mental states with a specific causal role in our practical reasoning. They are inputs to our actions since

they direct our conduct and prompt us to act according to their content (Bratman, 1999a, p. 18). Michael Bratman defines them as “conduct-controlling pro-attitudes” that function as inputs to further reasoning and, ultimately, as inputs to action (Bratman, 1999a).

Bratman (1999b) notes that rational agents can set up future-directed plans whose function is to structure and influence their conduct so that they can work on tasks or projects over the course of time. Intentions are the components of plans. Plans are, then, composite mental states that an agent has when she settles in advance on a specific course of action by designing the steps and stages that enable her to perform that action in the future. This is how agents successfully achieve complex goals they want to bring about (Bratman, 1983).

Plans are formed by different kinds of intentions organised hierarchically. These are what Bratman calls intentions concerning ends or intentions concerning means (Bratman, 2014). Intentions concerning ends are general, plan-like intentions that have the goal(s) an agent wants to bring about as their content. Intentions concerning means are more specific intentions regarding the various means and steps to achieve the end canvassed by the more general intention. If I want to go on a solo trip to the mountains, for instance, ‘I intend to go on a solo trip to the mountains’ would be my intention concerning an end. Upon forming this intention, I come up with other intentions concerning means that would help me to act on the previous intention. Such intentions concerning means would specify the mountain I want to go to, the train I need to take to reach the starting point of my hike, when to buy the train tickets, what to pack and so forth. The combination of all these intentions concerning ends and means form the structure of a plan. Being a more general kind of intentions, intentions concerning ends subsume intentions concerning means, where the latter are specific types of intentions that relate to the more general one.

Bratman defines intentions and plans as conduct-controlling pro-attitudes that play a motivational role in our practical reasoning. Intentions and plans direct our conduct and prompt us to act in relevant ways (the ways we intended). But Bratman does not talk about intentions and plans as reasons for action. Even though the motivational roles they play in our mental economy resemble those of reasons, he does not want to argue that they *are* reasons. The debate on whether intentions are reasons for action is vast and complex.²¹ In the previous chapters, I referred to intentions and plans as those things that provide agents with reasons. Here is what I mean by that.

Agents often deliberate on what to do and what their balance of reasons requires. They have a reason to do something, which might either justify their actions, motivate them to act or both. Agents then can form an intention to do something when they settle the matter on a debate and commit to a specific course of action. That intention can facilitate the agent's response to the reason the agent has because intentions, as commitments to actions, direct our conduct towards that action and prompt us to take the relevant steps to so act. Similarly, for complex actions that require agents to act over time, agents can make plans that would facilitate their response to reasons. More on this later in the chapter. In saying that the agents' intentions and plans provide them with reasons is not to say that intentions and plans give agents new reasons that add to their balance of reasons. That intentions and plans provide reasons means that they establish a connection in agents' mental economy between the reasons they have and the actions that will allow them to respond to those reasons. If I form an intention to walk the dog, that intention does not give me a new reason to walk that dog that I did not have before. If that happened, I would have two reasons to walk the dog (the one I already had, and the one given to me by my newly formed intention) which would tip my balance of reasons in favour of that particular action. And we want to avoid saying that forming intentions allows us to bootstrap

²¹ See Broome (2001); Brunero (2007) and Yaffe (1995) for some relevant discussions on this debate.

any sort of reason into existence. However, if I have a reason to walk the dog (motivating or normative, sometimes both), I might intend to walk the dog as a result of that reason, where that intention commits me to act that way. When I finally walk the dog that intention allows me to respond to a reason I had.

Intentions and plans connect agents with reasons they have. Plans and intentions can organise an agent's conduct and facilitate the process of responding to the reasons that prompted the agent to intend or plan in the first place. As conduct-controlling pro-attitudes, intentions and plans reflect the reasons agents have because they are formed to organise conduct in such a way that is conducive to acting on that reason. Intentions and plans justify, motivate or explain our actions to the extent that they reflect reasons that justify, motivate or explain someone's conduct. This is what I mean when I say that intentions and plans provide agents with reasons or give reasons. I often say that the reasons plans and intentions provide are motivating reasons when it comes to what the agent's reason for acting was. These reasons might also be normative if they were reasons for the agent to do something (if they justify the agent's actions). However, it is worth emphasising that the normative and motivating aspects might not always go together, and they can come apart.

Back to the elements of planning agency now. There are specific norms of rationality that govern the relations between intentions concerning ends, intentions concerning means and the agent's beliefs. Bratman (2014) inquiries into two norms, consistency and means-end coherence. For a plan to rationally prompt agents to act to bring about (complex, future) goals, the intentions that form it need to be consistent with each other and with the agents' beliefs. If an agent believes that her car has broken down, she could (or should) not form a plan to drive her friend to the airport with her vehicle. That plan would be inconsistent with her belief that her car has broken down and, as a consequence, it would result in her not being able to discharge the plan. Similarly, intentions concerning means need to be coherent with intentions

concerning ends, that is, I should form intentions concerning means that contain details I believe will lead me to successfully execute my end. To return to a previous example, if I have an intention whose end is to go on a solo trip to the mountains, I should not form an intention to wear high heels as part of my outfit if I believe that the means of hiking in high heels would not lead me to successfully execute my end to go up the mountain. These two norms of practical rationality are fundamental because intentions (and plans) would not succeed in guiding conduct if they were inconsistent and if intentions concerning ends were not coherent with more specific intentions concerning means.

Following Bratman more closely, we can see that there is an additional norm of practical rationality that governs the functioning of planning agency. According to this norm, agents must treat the reasons provided by plans and intentions as exclusionary (at least initially). Just as intentions would not succeed in guiding conduct if they were not consistent and coherent with one another and with agents' beliefs, plans (and their related reasons) would not effectively guide conduct if they were not treated as exclusionary. The idea is that, upon forming a plan, agents have to treat plan-related reasons as exclusionary for that plan to successfully guide their conduct. The default position for each plan formed is that agents do not reconsider it. They can reopen deliberation on the plan if needed, such as when there is overwhelming evidence that the plan is bad, when the agent's beliefs change and so forth.

To see why it is the case that agents must treat plan-related reasons as exclusionary, I need to go back to Bratman's account. Bratman does not say that the reasons that stem from individual plans need to be treated as exclusionary. However, there are two senses in which his account hints at this possibility. Firstly, Bratman (1999a) argues that, once formed, plans need to be resistant to reconsideration. They would not effectively serve their purposes of organising and coordinating conduct if they were changed at every turn. In this sense, they exclude reasoning in ways that go against the plan. Secondly, plans restrict the range of options

available to the present self, and that is a mark of what we call ‘exclusionary reason’ (Raz, 1986). In forming related sub-plans, an agent must eliminate those alternatives or options that are not in accord with the general plan. Plans can be of course modified if the circumstances change or if more stringent reasons to do otherwise arise.

It is evident, then, that Bratman assigns exclusionary functions to individual plans. Exclusionary means to cut off, to halt deliberation: this is a mechanism that prevents agents from deliberating further and stops them from reconsidering a certain reason or action one has settled on. In arguing that plans need to be resistant to reconsideration and that they restrict the options available to the agent (Bratman, 1999a), Bratman is attributing exclusionary functions to individual plans.²² Agents need to deliberate and weigh reasons for action before doing something, but there is a limit to the time they can actively spend deliberating. Unlimited deliberation might lead to inaction, for considering too many options would make it difficult for the agent to settle on something and act upon it. Some decisions will certainly require more time, but individuals will eventually reach a point where it is preferable to stop considering options and just decide what to do. By restricting the options available to the agent and by actively preventing her from reopening the settled plans for reconsideration, individual plans enable the agent to reach present and future goals. Restricting options and halting deliberation make a plan exclusionary.

I have argued that agents must treat the reasons provided by plans as exclusionary. Doing so would prevent them from deliberating and deciding to do otherwise once they have settled on a specific plan. I also argued that treating reasons as exclusionary is a norm of practical rationality. But how would this work in practice? Suppose I have a reason to write a paper for the XYZ Conference. To respond to that reason, I make a plan that will guide my conduct and

²² Note that Bratman does not use the terminology of ‘exclusionary’ or ‘exclusionary reasons’.

eventually lead me to fulfil my goal to write the paper. Once I settle on that plan, I treat that plan and the reason that plan reflects as exclusionary (at least initially). This has some consequences for my practical reasoning. Treating that reason as exclusionary will prevent me from changing my mind at every turn so that I can stick to my original plan and write the paper for the conference. This treatment of my plan-related reason is essential for me to reach my goal. If I did not stop deliberating and reconsidering my decisions at every turn, it would be difficult for me to refrain from doing something else, where that would jeopardise the success of my plan. Furthermore, treating that reason as exclusionary restricts the range of options available for me to choose from once that plan is in place. My plan to write a paper for the XZY Conference, when I treat it as exclusionary, prevents me from accepting my friend's invitation to dinner. I know that, if I go out to have dinner with my friend, I would not be able to finish the paper on time. Treating my reason to write the paper as exclusionary, would filter out options that are, in my eyes, incompatible with my plan, thus enabling me to reach my goal.

Treating certain reasons as exclusionary, then, is important to ensure the success of our plans. However, we also have the freedom to reopen our deliberation and change our minds if necessary. If more stringent reasons to do otherwise arise, and if it is not possible or rational to follow the original plan anymore, we can stop treating a reason as exclusionary, reopen deliberation on what to do, make a new plan and act on different reasons.

Before moving on to some explanatory models of planning agency, I need to clarify certain aspects of reconsideration of (past) plans. Refraining from reconsidering past decisions or plans does not mean that agents have to rigidly stick to those decisions or plans at all costs. What I mean when I say that agents must treat plan-related reasons as exclusionary for the plans to successfully guide their conduct is that it is appropriate for agents not to reconsider a plan they settled on. It is acceptable and rational to go along with a plan made in the past and exclude some options that are incompatible with it for the sake of facilitating the fulfilment of

goals and saving time on decision-making. This initial stance, not reconsidering a plan one settled on, can be defeated and there are cases when it is appropriate to reopen deliberation, reconsider a past decision and change one's mind. Among the clear cases when it is appropriate for an agent to reconsider a previous plan, we have those cases when our initial deliberation was flawed. We might change our beliefs about something and, when that happens, we realise that one (or more) of the premises used in making a plan was false. At that point, we have to reconsider, or we would be acting on false beliefs. Another case that involves flawed reasoning in formulating a plan is when we forget to consider pre-existent commitments (e.g., that we promised or agreed to do something) that might affect the discharging of our plans. In our previous example, if I promise my friend to go to dinner with her on a specific day, I need to factor in this commitment in my schedule and work around that. If I forget about this commitment and remember it when there is a clash with one of my plans or sub-plans, it would be permissible to readjust the plan to accommodate this commitment I had not considered whilst planning.

There are also cases in which we can reconsider a plan that is not based on false premises. These cases involve a change in preferences or values. These changes warrant reconsideration of past plans to accommodate new preferences or values we embraced. For instance, if I have a plan to cook steak every Sunday for dinner, I can reopen deliberation and modify that plan if, as a result of a value change, I go vegan. Another clear case when it is permitted to reconsider is when more stringent, reasons to do otherwise unexpectedly arise. If, in following my plan to submit a paper for the conference, I am writing but witness a hit-and-run out of my window, I now have a reason to go help the person hit by the car, where that reason would be weightier than the one to finish my paper. These are, then, some clear cases when it is permissible to reconsider a past plan. This is not an exhaustive list, and more cases will fit the bill for appropriate reconsideration. The main takeaway is that the fact that we must treat plans and the

reasons they provide as exclusionary does not make practical reasoning rigid. Agents can reconsider their past decisions and plans in a variety of cases and this allows for flexibility in decision-making processes.

That agents can reopen deliberation and effectively stop treating plans as exclusionary when needed, makes this Bratman-style notion of exclusionary reasons preferable to Raz's. Treating a plan as exclusionary speeds up our decision-making and allows us to fulfil our goals. If we could not reconsider our plans, though, we would risk having to act on false beliefs (when we know they are false), ignore pre-existing commitments we did not account for in planning and being unable to act on changed values or preferences. Being able to change one's mind and plans allows for flexibility in practical reasoning and is important for our autonomy, as I discuss in section 1.3 below. In Chapter 5, I reflect on reconsideration in the context of authority's plans to allow for the kind of flexibility we have in individual planning agency.

3.2 Explanatory models of planning and plan-related reasons

In this section, I canvass three explanatory models that depict the psychological processes behind planning that lead up to our actions. I also dive deeper into plan-related reasons and how plans reconnect agents with reasons for actions they have. I conclude the section with an account of individual practical reasoning that ties all these elements together.

The three explanatory models of what goes on when we enact plans are the belief-belief model, the belief-desire model and the belief-plan model.

Somewhat formally, the belief-belief model can be rendered this way:

$$B(P \varphi \text{ at } t) + B(\text{now is } t) \rightarrow \varphi\text{-ing}$$

According to this first model, what leads you to action is the belief that you have a plan P to ϕ at time t coupled with the belief that the time to enact your plan has come (i.e., the belief that now is t). The first belief in this model presupposes acceptance of the plan in question. When you create a plan, acceptance of that plan is crucial to carry it out. Without acceptance of the plan, you would not identify with it and would not recognise it as your own, with the result that the plan would not motivate you to action. In the individual case, when an agent acts alone and creates a plan, acceptance of that plan is straightforward: the fact that the plan is created by you makes your acceptance of it automatic. In this model, acceptance is implicit in the belief that you have a plan to ϕ at t . In practice, I form (and accept) a plan on Monday to cook dinner on Tuesday at 7 pm. Suppose that now it is Tuesday at 7 pm. I have a belief that I created (and accepted) a plan to cook dinner at t , and I also believe that now is t . These two beliefs combined lead me to act on my plan to cook dinner now (Tuesday at 7 pm).

The second model (somewhat) formally goes like this:

$$B (P \phi \text{ at } t) + B (\text{now is } t) + D [(\phi \text{ at } t) \text{ if } (P \phi \text{ at } t)] \rightarrow \phi\text{-ing}$$

In this case, there is a belief-desire combination that leads you to action. As in the first model, you believe that you have a plan to ϕ at t and you believe that now is t . These beliefs are accompanied by a desire to ϕ at t if there is a plan in place to ϕ at t . When you have these beliefs and desire you are prompted to ϕ at t . In this case, acceptance of the plan is embedded in the desire to ϕ at t . Again, I have a plan to cook dinner on Tuesday at 7 pm. Now it is Tuesday at 7 pm. I believe I have a plan to cook dinner at t , I believe that now is t and I have a desire (which signals my acceptance of the plan) to cook dinner at t that is conditional on having a plan to cook dinner at t . This belief-desire combination leads me to act on my plan to cook dinner on Tuesday at 7 pm.

Finally, the last explanatory model:

$$B(\text{now is } t) + P(\varphi \text{ at } t) \rightarrow \varphi\text{-ing}$$

This third model explains planned action by relying on the combination of a belief and a plan. There is a belief that now is t and a plan (made in the past) to φ at t . The belief that the time has now come leads to action because there is a pre-existing plan in place to φ at that particular time. Acceptance of the plan is, again, implicit for the agent who created the plan. In this case, I would form (and accept) the plan to cook dinner on Tuesday at 7 pm. If I believe that now it is Tuesday at 7 pm, that belief and my pre-existing plan would prompt me to act accordingly.

The difference between the belief-belief model and the belief-desire model hinges on two different ways of explaining intentional action. We might think that we do not need desires to be motivated to act and that beliefs, together with intentions and plans, are sufficient to motivate agents to act. If this is the case, we might prefer the belief-belief model to explain intentional action. The Humean theory of motivation (Railton, 1986; Smith, 1994), on the other hand, explains intentional action by appealing to both beliefs and desires. According to this theory, beliefs and desires are both needed to motivate the agent to act, where beliefs are mental states designed to match the external world, and desires are mental states that might or might not match the external world (they can indicate how we would like the world to be) but motivate agents to act. Theorists who espouse the Humean theory of motivation, then, would appeal to the belief-desire model to explain intentional action. Readers can choose their preferred way of explaining intentional action based on the intuitions they have on the matter.

I now develop my account of the relationship between plans and reasons. In section 1.1, I argued that plans give agents reasons for action. I also argued that these plan-related reasons are motivating reasons that agents have to treat as exclusionary, where this exclusionary treatment of plan-related reasons is a norm of practical rationality that ensures the proper functioning of plans. A motivating reason is something individuals base their actions on,

something they take as a motivation for their actions. We might wonder how agents act on those motivating reasons when they follow plans made in the past. After they form a plan in the past for future action, agents simply act on the plan when the established time comes. They, then, base their actions on the pre-existing plan. In which sense can we say that, when we have a plan, we are still basing our actions on a motivating reason? In other words, we want to know how following a plan allows us to act on motivating reasons that we must treat as exclusionary.

We base our actions on motivating reasons. This means that our actions are motivated by certain reasons and that we can invoke those reasons when someone asks what prompted us to act in a certain way. We can also base our actions on something when we have a mental attitude (i.e., mental state) towards it. This mental attitude can be a belief, a desire, an intention and even a plan. In this case, we can point to a belief, a desire or a plan of ours when someone asks what prompted us to act in a certain way. Let us suppose we have a plan to do something.²³ In this case, we base some of our actions on that plan. This means that the plan will motivate us to act according to its content and that we can refer to it in deciding how to proceed in enacting it. That plan might also explain why we are acting in a certain way, an explanation that we give to ourselves and others. If someone asks why we are doing something, or if we ask ourselves the same question, our plan might serve as an explanation for our conduct: we are ϕ -ing because we have a plan in place according to which we ϕ in certain contexts or at a designated time and so forth. As to how plans and mental states, in general, lead us to act on them, that will depend on the explanatory model we choose. We may be led to action by a combination of two beliefs (first explanatory model), by a combination of beliefs and desires (second model), or by a belief and a plan (third model).

²³ This can be a plan we immediately carry out in the present, or it can be a plan we make in the past to do something at a future time.

We have seen, then, how we base our actions on mental states. When we enact plans, though, we might think that we are acting without a reason or that the reason we are acting on is (merely) that we have a plan. This is because, when we enact a plan, we are doing what the content of the plan requires us to do, and it appears that our (previous) plan is the only element on which we are basing our action. However, this is not the case. Plans, as composite mental states, are ‘vessels’ that contain the reasons they were created for. In this light, plans play two specific roles in our practical reasoning. They help us to respond to reasons and, when we enact them in the future, they reconnect us with the reason(s) we had for creating them.

Firstly, plans might help us to better conform to reasons. We have reasons to act in certain ways. Often, organising and coordinating our conduct help us to do what we have a reason to do. We best organise and coordinate our conduct via plans. Therefore, plans help us to better respond to our reasons. Often, we have a reason to act in a certain way, where the action in question is a complex activity that requires careful organisation and coordination of our thoughts and actions. Go back to my reason to write a paper for the XYZ conference, where writing a paper is a complex action that stretches over time. Organising and coordinating my intentions and conduct would help me respond to my reason. The best way to organise and coordinate my conduct is to plan it. Therefore, a plan to write a paper for the XYZ conference would help me to better respond to the reason I have to write that paper.

Secondly, plans reconnect us with the reasons we formed them for. If we have a reason to do something, we can create a plan that will help us to respond to that reason. In the previous example, my plan to write a paper will allow me to respond to the reason I have to write that paper. Perhaps, I will lose my funding if I do not write this paper, so the reason that motivates me to write it is that I cannot lose my funding. Forming a plan with different steps (e.g., research the topic today, create a paper structure tomorrow, start writing the day after tomorrow and so forth) allows me to make progress towards the completion of the paper and organises my

conduct to complete this task. When I follow my plan and start writing the paper, am I still motivated by the reason not to lose my funding I considered three days ago? Yes, I am still motivated now by the reason I considered three days ago because that reason prompted me to form a plan in the first place. If I had not been motivated by the reason not to lose my funding, or if I had not had a reason to write in general, I would not have formed a plan to write the paper. This plan exists because of my motivating reason not to lose my funding. There is then a causal relation between my reason not to lose funding and the plan to write the paper: that reason caused me to plan accordingly. I want to say that this causal connection between reasons and plans makes it so that plans are metaphorical vessels that contain the reasons they were created for. My plan to write the conference paper, created for the reason that I cannot lose funding, contains that reason which I can act on in the present or the future when I discharge the plan.

This particularly matters when we make plans in the past that we follow far in the future. If I am motivated by a reason not to lose funding in February and I plan accordingly, I might not follow that plan immediately. I might have a deadline for that paper in June and act on my plan to write in June. Since the plan was created because of that reason, we can say that, when I act on my plan to write the paper in June, I am acting on the reason not to lose funding that motivated me to create that plan in February.

So, this is how plans (re)connect us with the reasons they are based on. By following the plan, we respond to the initial reason that prompted us to create the plan.

To connect us back with a point made earlier, agents must treat these plans and plan-related reasons as exclusionary. We can create plans to organise our conduct and respond to reasons. Once formed and executed, plans connect us with the reasons they are based on but they have to be treated as exclusionary, where this exclusionary treatment of plans (and the

reason they were created for) is a norm of practical rationality for the correct functioning of plans.

In practice, I might, on Monday, have a reason to cook dinner on Tuesday at 5 pm and, as a result, I come up with a plan to help me act on that reason. The plan would not effectively guide my conduct if it was not treated as exclusionary. Deliberating at every turn and frequently changing my mind regarding what to do, would not successfully lead me to follow my reason to cook dinner at that designated time and day. Because I care about my reasons and I want my plan to succeed, I must treat the first-order reason the plan gives me as exclusionary. That will prevent me from deliberating and reconsidering my plan unless necessary.

Suppose the time has now come, and I rightly believe that it is, in fact, Tuesday at 5 pm. I, then, remember I have a pre-existent plan to cook dinner at this time today (Tuesday) and I proceed to act on it. In enacting this plan I made in the past, I am relying on the plan for guidance on what to do but, through the plan, following my original reason to cook dinner at that time and day. I am not, however, reopening deliberation or questioning the original reason I based my plan on. This is because treating a reason (and a plan) as exclusionary, makes me simply follow the plan without reconsidering the deliberative process that led me to the creation of the plan itself. On Tuesday at 5 pm (or at any given moment), I mentally ‘scan’ my mental states for reasons and I come across the fact that I have a plan for that specific time and day. In scanning my mental states, I am thinking and deliberating but I stop my deliberation when I remember my previous plan. I do not reconsider or make changes to the plan, thus treating it as exclusionary.

Before moving on to section 1.3, I sketch an account of individual practical reasoning that includes all the components of practical reasoning I explained in the last two sections. This

account of individual practical reasoning constitutes the first stage of the three-stage model of practical reasoning to investigate the compatibility between autonomy and authority.

Partially following Bratman (1999a, 1999b), I sketch a five-step account of individual practical reasoning:

- 1) An agent, call her A, has a goal to achieve and a reason to do something.
- 2) A forms an end-related intention toward the goal and/or reason.
- 3) A forms means-related intentions connected to the previous end-related intention.
- 4) A has a plan to act on (in the present or the future).
- 5) By default, A must treat that plan (and the related reason in 1) as exclusionary.

The account summarises the previous discussion on individual practical reasoning. When agents have goals to achieve or reasons to do something, they usually form intentions concerning ends toward those goals or reasons. Agents then proceed to form related intentions concerning means to fill in the prior end-related intention (to fill in the prior plan) and to take the relevant steps and sub-steps to carry out the plan. The combination of a main end-related intention and its means-related intentions constitutes a plan on which individuals can act in the present or the future. Once agents have settled on a plan, they must treat that plan (and the reason the plan was created) as exclusionary. This is to ensure the plan can successfully guide their conduct. If required and necessary, agents can choose to reopen deliberation and change their mind on what to do.

3.3 Exclusionary reasons, autonomy and personal identity

This section aims to investigate whether the norm of practical rationality according to which plans and the reasons they were created must be treated as exclusionary impairs individual

autonomy. To do that, it also analyses some views of personal identity that would support the planning mechanisms explained in sections 1.1 and 1.2. This section starts with a concern that treating plans and plan-related reasons as exclusionary violates personal autonomy and ends with a brief analysis of diachronic planning agency (i.e., planning agency over time). This concludes the analysis of Stage 1 of the three-stage model to solve the autonomy – shared agency – problem.

Treating plans and the reasons they were created as exclusionary has two main consequences in agents' practical reasoning. Firstly, it forces agents to stop deliberating and reconsidering their choices, thus interrupting their deliberative processes. Secondly, it reduces the range of options for agents to choose from in deciding what to do. We might think that treating reasons we based our decision on in the past as exclusionary poses a problem for personal autonomy. Does this norm of practical rationality that prevents us from doing certain things (i.e., because they go against the plan) compromise autonomy? A plan we made in the past effectively blocks our deliberation and filters out options that are incompatible with the plan itself. This seems to leave less room for deliberation for our future selves, who are constrained by choices made in the past. I argue that past plans do not impair future selves' autonomy. They facilitate individual autonomy by enhancing agents' reason-responsiveness capacity. I show how by analysing two different views on personal identity.

On the face of it, being led by exclusionary reasons to stop deliberating and reconsidering our choices negatively impacts our deliberative processes. In determining how to act, we want to evaluate all the relevant aspects of the case at hand and gather as much information as possible on all the options that are available to us. This is valuable in our practical reasoning and ensures that we act rationally and according to reason. However, carrying on deliberation indefinitely and frequently changing our minds might lead to inaction. If we were to consider all the pros and cons of every course of action accessible to us and compare it against a virtually

endless list of possible courses of action before making every decision, it would be very difficult for us to do anything at all. We are often required to promptly make decisions in our daily lives. We need a mechanism that facilitates our decision-making processes and helps us in making decisions. Plans are a way for us to settle the matter of what to do because they organise and coordinate our behaviour. Creating a plan and settling the matter on what to do, then, prevent inaction by helping us to choose a course of action. Choosing what to do, though, is not enough to prevent inaction. If we reconsidered previous choices at every turn, our plans would fall through, and we would not be able to successfully reach our goals. Treating plans as exclusionary prevents us from constantly reconsidering our past decisions and facilitates acting on past plans.

Treating plans as exclusionary facilitates our decision-making processes, it does not hinder them. However, treating plans and plan-related reasons as exclusionary reduces the range of options available for us to choose from in the future. In forming related sub-plans and means-related intentions at some time t , agents must eliminate those alternatives or options that are not in accord with the general plan and end-related intention. This means that the plans we make at t influence our thoughts and actions at t_1 . If we want to carry out our past plan at t_1 , we need to make some sacrifices and avoid things we might want to do because they are not in tune with the plan made at t . This seems to impair our autonomy. Previously I said that autonomy, as the exercise of self-governance, amounts to responding to reasons and deciding for oneself how to act. If the plans we treat as exclusionary dictate our present and future behaviour, take away some options from us and actively influence our actions, how can we decide autonomously? If we are bound by plans made in the past, which we treat as exclusionary, we cannot decide for ourselves how to act in the present, or so it seems.

Two points can be made in response to that. Firstly, being autonomous does not necessarily entail that we constantly evaluate our options or decide for ourselves what to do at

every turn. That would support a view on autonomy that cannot realistically obtain. Agents often have a limited amount of time to make decisions and act on them. Constant deliberation and reconsideration, as explained above, lead to inaction and do not enable agents to respond to reasons and act on them, which is the whole point of being autonomous. Secondly, plans can help us to respond to reasons and act on them. Plans made in the past indeed restrict the range of options available to the present self. However, having a filtered set of options is helpful in quickly deciding what to do and helps us to better respond to reasons. As we saw in section 1.2, we have a reason to do something. Organising and coordinating our conduct help us to do what we have a reason to do. We best organise and coordinate our conduct via plans. From this, it follows that plans help us to better respond to reasons. Autonomy is, among other things, a matter of reason-responsiveness. Treating plans as exclusionary does not prevent agents from exercising their autonomy precisely because plans help agents to better follow reasons and make decisions, thus boosting rather than impairing their autonomy.

Plans and the reasons they were created for, which we treat as exclusionary, do not threaten individual autonomy. And it is important to insist on the fact that plans made in the past influence and direct our present and future conduct. Plans made in the past stretch over time and reconnect our present and future selves with reasons for action we based our decisions on earlier.²⁴ Plans are vessels that (re)connect us to previous reasons for action. When we act at t_1 on plans we made at t_0 , we are acting at a distance on a plan we created before (at t_0). This is how reasons influence us over time. We do not need to act immediately on reasons for action that apply to us. We can make plans that enable us to act on reasons at a later time,

²⁴ We can also say that the diachronic agent might be more reason-responsive by having one temporal part look at one set of reasons, and another temporal part at another, thus effectively dividing labour between temporal parts.

especially when reasons are associated with complex goals that require various steps and effort to be achieved.

We might think that planning our conduct over time (diachronic planning agency, Bratman, 2018) and reconnecting our present self with past reasons via plans works only if we espouse a relatively unitary notion of self.

Bratman (2018) argues that to diachronically govern our life we need to be the same person who is always doing the governing (p. 9). Bratman identifies two necessary conditions for diachronic self-governance: a) being the same person over time, and; b) having a web of interrelated intentions and plans. If we accept this unitary picture of diachronic planning agency, we might find it easy to see how agents can make plans in the past that they will act on later in the future. If by the sameness of person over time, we mean that there is a unitary stream of consciousness where all mental states are interconnected and tied to the same self, we could argue that the self can have complex goals, think of its future and remember its past intentions and plans. The self will be able to connect thoughts and actions to itself and to recognise them as its own, thus ensuring those continuities and connections that are necessary for planning agency over time. And we might think that, in this case, there is no doubt that plans made in the past, which we treat as exclusionary, do not hinder individual autonomy. It is, after all, the same self who made a plan at t_0 that is now enacting it at t_1 . This unitary, conception of the self seems to pose no threats to autonomy when there is one stream of consciousness all intentions, plans and action are tied to.

However, one might endorse a different conception of the self that is not as unified. Some accounts of personal identity, for instance, see the self as a succession of separate time-slices, with no underlying person or unitary sense of self that accompanies the time-slices (Parfit,

1971).²⁵ This account of the self raises some questions. If we endorsed this less unitary notion of the self, would we be able to plan our conduct in the same way? And would we see plans made by a past time-slice as ‘our plans’?

To answer these questions, let us briefly examine David Lewis’ account of personal identity. Lewis (1983) argues that a continuant person, namely, a person that is the same at various stages in time, is the sum of several person-stages, each of which is related to all the other stages and to itself through ties of mental continuity and connectedness, and each of which is not a part of another continuant person (p. 6). Two main relations matter to Lewis regarding personal identity over time. The first of these is the relation of mental connectedness and continuity that occurs among person-stages. Connectedness refers to the similarity and causal connections that we have between present and future mental states: changes between mental states that belong to time-slices of the same person should be gradual and causally dependent on one another. Continuity is described by Lewis as “the existence of step-by-step paths from here to there” (1983, p.2), which are paths that bring us from one mental state to the other and emphasise their interrelatedness. Continuity and connectedness form a relation among person-stages or time slices. The second relation that matters here is identity itself, which occurs when a continuant person, with temporally extended time-slices of itself, is the same over time. When several mental states that belong to different time-slices share a relation of mental continuity and connectedness, and when different time-slices constitute an aggregate of interrelated mental states which do not refer to other aggregates, we can talk about a person whose identity survives over time.

²⁵ In this debate about different accounts of personal identity, there is an interesting response to Parfit from Christine Korsgaard. In this reply, she argues that many of our actions are intrinsically diachronic, not just in the sense that we need to cooperate with ‘future selves’ to get them, but that even their realization takes time (Korsgaard, 1989).

Now that we have presented Lewis' theory of personal identity, we can go back to the two questions that matter to us. Is it possible, when we endorse a theory of this sort, to plan our conduct in the way we described in 1.1 and 1.2? And would we consider plans made by a past time-slice *our* plans? The first question is about the creation of plans by a past time-slice of ourselves and how these plans are acted on by our future time-slices.

We have said that agents can have end-related intentions to respond to their reasons and correspondent means-related intentions that together form a plan. End-related and means-related intentions, as mental states that are part of the same plan, need to be consistent and coherent with one another for them to successfully prompt agents to action. The mental states that constitute plans, then, already stand in a relation of connectedness between one another that brings time-slices back to previous choices made by past time-slices of the same person. If an agent has a means-related intention to buy pasta from the local shop, for instance, that intention is connected to the general, end-related intention to cook Italian for dinner. The two intentions need to be consistent and coherent with one another to form a proper plan and, normally, when an agent thinks about one of those intentions, her attention is immediately drawn to the other intention that is part of the same plan. If the means-related intention to buy pasta from the local shop has to be enacted at t_2 , when t_2 comes, the t_2 time-slice will be reminded of the plan upon remembering that she needs to go to the local shop to fetch the pasta. The interconnections between mental states of the same plan ensure that agents are brought back to the plan that a past time-slice of them made at t_0 whenever they need to act on a means-related intention. This means that complex plans often stretch far into the future, and their interconnections ensure that different time-slices of a person can take care of an aspect of a plan and, eventually, bring it to completion. Present time-slices can, of course, forget about the existence of a plan made by a past time-slice, or decide not to carry out the plan. Nevertheless, the interconnections between the mental states that are part of the same plan, and the reasons

those mental states provide, can influence agents' actions over time, where those interconnections ensure the agent's time-slices are reminded of every part of the plan at several time-points.

Via plans, which contain the reasons they were created for, future time-slices of the self have access to all the past reasons of previous time-slices, and can follow them in the future. Acting on a plan made by a past time-slice, then, means following reasons past time-slices of ourselves had and bringing about complex goals that require related actions overtime. This enhances the reason-responsiveness of future time-slices and facilitates their autonomous decisions and actions. As an entity with a limited cognitive ability, a time-slice cannot do all the work necessary to figure out what to do to bring about complex action. And it does not have to. It can rely on past time-slices' decisions and follow their plans. This would enhance its reason-responsiveness (and the autonomy of the overall self as a sequence of time-slices).

Regarding the second question, we want to know whether a present time-slice can identify with plans made by another, past time-slice and recognise them as its own. The answer to this question lies in the relations of continuity and connectedness among time-slices Lewis talks about (Lewis, 1983). If several time-slices are related to one another via causal relations of continuity and connectedness, and if the sum of all those interconnected time-slices is associated with only one person and no one else, it is easy to see how future time-slices can identify with plans made by past time-slices and call them their own. Continuity and connectedness among time-slices ensure that the transitions and changes that take place between two or more time-slices are generally gradual and not abrupt. If a past time-slice forms end-related and means-related intentions to be carried out at different times in the future, the close relations between stages of the same person make it likely that future time-slices remember the plan and are related to it closely enough for it to be carried out. The different mental states that compose a plan, as we have said, are also interconnected, and this helps a

person-stage remember the plan and associate it to a past stage. Suppose we have a timeline that goes from t_0 to t_5 , where a time-slice of a person occupies different points in time. Suppose an end-related intention and several means-related intentions are formed at t_0 for future time-slice to act on them at $t_1, 2, 3, 4$ and 5 . Furthermore, the end and means-related intentions are coherent and consistent with one another, and there are relations of continuity and connectedness among the time-slices of a person. When all of this occurs, the t_5 time-slice carries out its assigned means-related intentions fully knowing that, as a time-slice, it has relations with the past time-slices, and that the means-related intention she is presented with is part of a complex plan conceived by the t_0 time-slice. When relations between a plan's parts and relations between time-slices are present, and when those time-slices form one and only one person, present or future time-slices can identify with plans created by past time-slices. This identification also supports the previous point that past plans and their constraints do not hinder personal autonomy if we conceive of plans as instruments through which individuals can respond to reasons and fulfil complex goals.

No matter the view of personal identity we endorse, then, we can see that plans made by past time-slices connect future time-slices with past reasons for action. This is fundamental to bring about complex goals that require a thorough organisation of thought and action overtime. Plans give future time-slices access to the reasons of past time-slices, thus enhancing their reason-responsiveness as a result. This helps future time-slices with their deliberation and does not impair their autonomy.

3.4 Summary

This concludes our analysis of Stage 1 of the three-stage model of practical reasoning that solves the autonomy - shared agency problem. In this chapter, we focused on individual

practical reasoning and analysed the mental processes behind individual deliberation and decision-making. I explained how agents intend and plan their conduct, as well as the norms of practical reasoning that govern the functioning of plans. Treating plans as exclusionary is a fundamental norm of practical rationality. If agents constantly reopened deliberation and did not take their plans as conclusive when appropriate, they would be victim of inaction and likely unable to fulfil their complex goals. However, treating plans as exclusionary, especially plans made in the past, considerably reduces the options available for agents to choose from and constrains further reasoning for them. In section 3.3, I argued that treating plans as exclusionary does not impair autonomy. On the contrary, it facilitates our decision-making processes by halting deliberation (at some point) and reducing the range of options for agents to choose from. Halting deliberation and reducing options speed up decision-making and prevent inaction. Ultimately, I made the point that plans help us respond to reasons by organising and coordinating our conduct. In Chapter 4, I analyse shared plans, that is, those plans created by more than one agent and establish whether they negatively impact our autonomy.

Chapter 4 - Shared Practical Reasoning, Autonomy and Shared Plans

4.0 Introduction

In Chapter 3, I argued that the apparent incompatibility between autonomy and authority is not an isolated phenomenon that only concerns the relationship between individuals and authorities or, in the case of political societies, the state and the governed. This tension between autonomy and authority is part of a more general tension between autonomy and shared agency.

Wolff sees obedience to authority as a matter of doing something because the authority tells you to do it (Bates, 1972, p. 176). When you are presented with an authoritative directive or a law, you should not deliberate on its merits or whether to follow it. You should simply obey it and act accordingly because the authority says so. In other words, you should treat that directive as exclusionary: the directive replaces your other reasons for acting otherwise and motivates you to act as the authority commands. When the state and the governed act together then, the state formulates plans of action for its governed in the form of laws or directives, and the governed must act on them without opening deliberation. This seems to threaten autonomy as the exercise of the capacity to deliberate and decide for oneself what to do. In Chapter 5, I argue that autonomy and authority can be compatible despite the governed having to treat authoritative directives as exclusionary. It is not immediately obvious, though, that that is the case and that is why Wolff's problem of autonomy and authority has plagued many political philosophers.

The problem is that plans in general, as I explained in Chapter 3, have an exclusionary nature. Once formed, they should not be reconsidered to successfully guide our conduct. In a way, we can say that we are doing something because there is a plan in place: we are doing

something because the plan tells us so. This is more evident when it comes to plans made in the past. When we settle on a course of action for the future and plan accordingly, we should not deliberate on the merits of the plan again when the time comes to act on it. We should follow it, treat it as exclusionary and act accordingly. This allows us to complete complex projects that require a fine-grained organisation of our thoughts and actions for a prolonged period. In Chapter 3, we saw that this is not a problem for individual practical reasoning when one plans and acts by oneself. The person who deliberated and planned is the same as the one who is now executing the plan. She planned to do something for a reason and following the plan without deliberating and reconsidering it allows her to respond to the reason that prompted her to create the plan in the first place. Treating plans as exclusionary in the individual case is unproblematic and does not negatively impact autonomy.

Things are more complicated when it comes to agents acting together. When multiple agents create a shared plan of action, different agents' choices, preferences and reasons feed into the creation of this plan. And it is not always the case that every agent involved in the shared activity contributes equally to the creation of the shared plan. When agents treat shared plans as exclusionary and act on them without reopening deliberation on what to do, they are accepting other agents' choices, preferences and reasons at face value. They are not deliberating and deciding for themselves how to act, which are the tenets of autonomy, as we established in Chapter 2. Does treating shared plans as exclusionary in this sense impair individual autonomy?

In this chapter, I analyse the relationship between autonomy and shared agency, by focusing on small, egalitarian contexts in which agents do things together without a hierarchical structure of authority. I show how autonomy and egalitarian (i.e., non-hierarchical) shared agency can be compatible. This chapter is divided into three main sections. Section 4.1 deals with the mental states involved in shared practical reasoning and the norms of practical

rationality that govern shared planning. Section 4.2 puts forward an account of shared practical reasoning that explains how those mental states and norms of practical rationality guide the conduct of agents involved in shared activities. Finally, section 4.3 explores the implications of egalitarian shared planning on individual autonomy. It shows that shared plans do not always hinder autonomy. The arguments of this section constitute the second stage of the three-stage model of practical reasoning to solve the autonomy-shared agency problem. Section 4.4 summarises the chapter's main points and arguments.

4.1 Joint Activity and Shared Plans

Agents often carry out shared activities and projects with others. Students work on joint presentations with other students, friends go on holiday together, cook together or go hiking together. What goes on in these cases? Agents might come together to devise a shared plan of action.

Take the case of a group of friends who want to go on holiday together and call it HOLIDAY. They might meet up to choose a holiday destination and plan their journey. Each of them might pitch ideas to the group and eventually, they will settle on a holiday destination. They then can decide who is going to take care of which organisational aspect of the trip, such as booking flights, finding accommodation and so forth. In addition to that, each of them must make time for the trip by organising her commitments in a way that does not disrupt this shared plan and is not incompatible with going away with friends. The friends cooperate to discharge the different organisational aspects of the trip by each doing their part and, eventually, they come to fulfil their original goal to go on holiday together. This is an example of shared activity that sees multiple individuals acting together and coordinating their actions to bring about a certain outcome. The theory of action has focused for a long time on joint actions performed

by two agents, (Searle, 1983; 1990; Tuomela and Miller, 1988), with classic examples such as two individuals walking together (Gilbert, 1996; 2000; 2015) or painting a house together (Bratman, 1992). Research on the topic has progressed to analyse cases of larger-scale joint activity with multiple individuals (Bazargan-Forward, 2022; Gilbert, 2006; 2014; Kutz, 2000; Ludwig, 2017; Shapiro, 2011; 2014), which will ultimately be our focus in talking about political societies as groups with a structure of authority.

Before starting our analysis of the mental states involved in acting together, it is important to clarify the terminology used in this chapter. Firstly, shared agency is used to indicate the phenomenon of acting together (Bratman, 2014), where agents cooperate with others to complete certain tasks. If shared agency indicates the general phenomenon of acting together, we can use shared activity or joint action (alternatively, joint activity) to refer to instances of agents acting together (i.e., specific cases of shared agency).

Secondly, an important distinction to be made is one between non-hierarchical and hierarchical shared agency. In understanding and researching shared agency, the literature initially focused on cases of acting together where individuals are on the same level and contribute equally to the creation and execution of a shared plan. This included cases such as two agents painting a house together (Bratman, 1999a, pp. 94-99), or singing a duet together (Bratman, 1999a, p. 94). However, more ordinary and complex examples of shared activity, show that agents are often not on the same level when they act together, where some of them are in a position of authority over the others. This asymmetry of power in acting together influences the way shared activities are planned, organised and carried out. For this reason, it is important to distinguish between non-hierarchical and hierarchical shared agency. Non-hierarchical shared agency, or egalitarian shared agency, deals with cases of individuals who are equals and act together in *ensembles* that lack a structure of authority. ‘Non-hierarchical’ does not mean that every participant in a shared activity is in the same situation as other

participants. There might be economic, social and cultural differences between agents in a shared activity. By non-hierarchical, I simply mean the absence of authority structures amongst participants, and the fact that no one is in a position of power over other agents.

Hierarchical shared agency comprises shared activities with a hierarchical structure of authority where one or some participants have authority over others who follow the authority's plans. This authoritative structure creates asymmetries of powers between agents or group members, with agents having different roles and responsibilities. This chapter deals with non-hierarchical shared agency and analyses cases of agents acting together without a structure of authority. Hierarchical shared agency is analysed in Chapter 5, where I delve deeper into the last stage of the three-stage model to solve the autonomy-shared agency problem.

It is now time to focus on the components of non-hierarchical (egalitarian) shared agency, that is, the mental states and norms of practical rationality that enable agents to act with others. I will limit my analysis to a handful of these mental states that I deem fundamental for me to build an operative account of egalitarian shared agency.

What are shared goals? In discussing individual practical reasoning in Chapter 3, I defined goals as the contents of an agent's final desires. Another way to think about goals is by seeing them as some properties of the world we value, properties that explain why we have a certain pattern of final desires. We might not be aware that we value these properties, but they can nonetheless shape our desires and actions. In acting with others, things get complicated as different agents might value different sets of properties that need to be somehow reconciled to produce joint action. Sometimes agents sit together to discuss what properties of the world they value and direct their actions towards them. I value helping students find meaningful career paths after college and I discuss this with two other people on LinkedIn who also value the same thing. We discuss ways in which we can help students together and direct our actions

towards them. Other times, individuals tag along in a shared activity that is undergoing and, by observing what other agents are doing, they can identify what outcome those agents' actions are directed to. I am at a friend's garden party, and I see a few people passing some plates along. I infer they are trying to set the table and, since I value helping, I start passing plates and cutlery to set the table. I thus direct my actions towards setting the table by inferring that setting the table was our goal. In the context of joint actions, we can define goals as the outcome an action is directed to (Butterfill and Apperly, 2013). This will require participants in a shared activity to recognise that the actions they are performing are directed towards bringing about a certain outcome. They do not need to have all the details of the outcome they want to bring about, but they need a general sense that what they are doing, what they are contributing to with their individual action, is directed at an outcome they are bringing about with others.

Agents can have intentions that are directed towards an outcome they would like to bring about with other agents. When acting with others, agents individually intend to do something with others, that is, they individually intend to do X with Y. For instance, I intend to set the table with my friends at the garden party, I intend to move the couch upstairs with my neighbour, I intend to cook dinner with my partner and so forth. Bratman calls these intentions "intentions that we J", where J is a joint action that multiple agents are carrying out or will carry out (Bratman, 2014, p. 13). Here I am referring to paradigmatic and simple cases of acting together, where it is straightforward for agents to decide what to do together and when agents acknowledge their actions are directed toward a joint outcome. Not all cases of joint action will be this simple and agents can act together without all of them intending to do X with Y. Hierarchical shared agency in large-scale groups will often be of this kind: not everyone will intend to do something with others. However, the authority's guidance will be able to

coordinate agents' actions towards certain outcomes. For now, let us focus on paradigmatic, simple cases of egalitarian shared activity.²⁶

'I intend to do X with Y' is an end-related intention whose content is the outcome (goal) the agents' actions are jointly directed to. When agents acknowledge that the actions they are performing or will perform in the future, are directed towards jointly bringing about an outcome, they form intentions to do something that would bring about that outcome with others. Often, when individuals act together as equals, a division of labour takes place whereby the shared activity is divided up into smaller tasks, and each participant executes one of those tasks. For instance, if we are cooking together and planning on having pasta and a side salad, I might be responsible for making the salad while you take care of making pasta. In this case, I have a general intention to cook with you, and a related intention to make salad, which is my task in the joint action of cooking together. Agents then can each have means-related intentions to carry out their portion of the shared activity and end-related intentions to act with others.

End-related intentions to do X with Y and means-related intentions to do one's portion of X form a plan that is shared by two or more agents. A plan is shared when it is devised to coordinate the conduct of multiple agents acting together to bring about specific outcomes (outcomes they are generally aware of in simple cases but might not be aware of). Shared plans are clusters of end-related and means-related intentions that structure and guide the conduct of participants in a joint activity by giving these participants reasons for action.

Shared plans give reasons in the same sense individual plans give reasons. In Chapter 3, I explained that an agent has a reason to do something and, as a result, creates a plan of action that would facilitate her response to that reason. This plan of action does not provide the agent

²⁶ To see how cooperation or joint action is possible even without sharing intentions (or having collective intentional mental states) see Ritchie (2020). For a taxonomy of different kinds of cooperative activities that range from full cooperation to minimally cooperate behaviour see Ludwig (2020).

with a new reason but merely reflects the reason the agent has to plan since the plan was created to respond to that particular reason. Shared plans work similarly. Agents have a reason to do something with others and they can plan their conduct to act on that reason. I have a reason to cook dinner with my partner, and my partner has a reason to cook dinner with me. Let us assume there is a complete overlap of reasons between the two of us, where we both have a reason to cook with the other. We make a plan with an appropriate division of labour that would allow me to act on my reason to cook dinner with him and allow him to act on his reason to cook dinner with me. In this case, the shared plan does not give us a new reason to do something but re-connects us both with the reasons that prompted us to plan with the other. Since shared plans often involve a division of labour among different parties, with several means-related intentions that fill in the original plan, each agent would also have reasons to carry out specific tasks that relate to the general plan. Since we are having pasta and a side salad, and since I am in charge of preparing the side salad, the plan gives me a reason to prepare a side salad. This reason does not add to my balance of reasons: it does not count as yet another consideration in favour of cooking with my partner. It is a constraint of the shared plan that reconnects me with my original reason to cook with my partner. If I had not had a reason to cook with my partner, I would not have formed a plan to cook with him and I would not have been subjected to the constraint of preparing a side salad.

In the example above where I intend to cook with my partner and my partner intends to cook with me, I assumed there was a convergence of preferences and reasons between me and my partner. We both have reasons to cook with one another and our intentions and shared plans are created to respond to those reasons. However, acting together is compatible with agents bringing about a common goal for different reasons. For instance, my friend and I are painting her house together. She is motivated by a reason to renovate her house (which includes painting), while I am motivated by a reason to repay her. She helped me a few months ago and

it is now my turn to do her a favour. I do not have a reason to paint her house and I am not invested in her renovation project. I simply want to repay the favour as I do not like being indebted to people. Despite being motivated by different reasons, I can still intend to paint the house with her, and she can intend to paint the house with me. The constraints that come from the plan's means-related intentions can guide our conduct even though we are motivated by different reasons.²⁷

The norms of practical rationality that govern individual planning agency (Chapter 3) apply to joint activity and planning too (Bratman, 2014). The norms of rationality that govern how intentions and plans operate are consistency, means-end coherence and exclusionary treatment of plans. Plans and intentions need to abide by these norms to effectively lead agents to action and guide their conduct. Regarding consistency, the intentions that constitute a shared plan, that is, the intentions of the participants in a shared activity, need to be consistent with one another and with the agents' beliefs for the shared plan to be effective. If two people have a shared plan to go out for dinner together at Restaurant X on Friday night, each of them will (most likely) have an intention directed towards the shared goal, an intention to go to dinner at X on Friday with the other person. These individual intentions need to be consistent with one another. If one of the two participants in this shared activity, for instance, had an intention to meet the other at Restaurant Y on Friday night, that intention would not be consistent with the other participants' intention to go to Restaurant X on Friday night, with the consequence that they would fail to meet up and the shared plan would not be executed.

Moreover, agents' intentions concerning means need to be coherent with their intentions concerning ends. Suppose that the two people who want to go to dinner together, A and B, each

²⁷ When agents act together to bring about desired outcomes, there is the question of whether they are entitled to the other agent's reasons for action, that is, whether they can act on another agent's reasons (Roth, 2017). I will explore entitlement to someone else's reasons for action, or transmission of reasons from one agent to other agents, in chapters 5 and 6.

have the intention to go to dinner with the other at Restaurant X on Friday night and suppose further that they decide that A will pick B up on her way to the restaurant at 6 pm on Friday. A's means-related intention to pick B up on her way to the restaurant at 6 pm on Friday is coherent with A's end-related intention to go to dinner with B at Restaurant X on Friday night. However, if A formed a means-related intention to drive straight to Restaurant X at 6 pm on Friday, that means-related intention would not be coherent with the end-related intention to go to dinner with B at Restaurant X because the shared plan involves A picking B up on her way to the restaurant. The means-related intention to drive straight to Restaurant X at 6 pm on Friday would enable A to successfully reach the restaurant, but it would not contribute to the discharging of the shared plan, as B would count on A to pick her up and would thus not be at the restaurant.

Furthermore, agents must treat the shared plans and their related intentions as exclusionary. When agents treat the shared plans as exclusionary, they stop deliberating regarding what to do at some point and do not reconsider the shared plan unless necessary. As a norm of practical rationality, treating plans as exclusionary is fundamental to their proper functioning. Plans are created to organise and coordinate thought and action. They would fail to do so if agents continuously reconsidered past plans and often changed their minds about what to do. When agents treat plans as exclusionary, they are largely prevented from reopening their deliberation, with the consequence that their plans will be a stable guide to their conduct and reasoning.

So, participants in shared activities must treat the shared plans as exclusionary. Doing so prevents them from deliberating and deciding to do otherwise once they have created a shared plan and agreed on a division of labour between participants to fulfil the plan itself. How does treating plans as exclusionary work in practice? Suppose that friends who want to go on holiday together (shared goal) create a shared plan that will help them reach that goal. They settle on a

destination and date for their trip and on the fact that each one of them will be in charge of a certain aspect of the plan (e.g., booking tickets, finding accommodation or creating an itinerary). Once participants in the shared activity have settled on the shared plan and its sub-plans, they must treat the plan and the reasons that it provides as exclusionary. This will prevent them from continuously changing their mind so that they can rationally follow the plan and bring it to completion. Furthermore, this exclusionary treatment of plans has other consequences for the participants' practical reasoning. Treating those reasons as exclusionary would not only prevent them from constantly changing their mind, but it would also restrict the range of options available to them to choose from. The shared plan of going on holiday with friends from the 1st to the 5th of July would render the option of going to a conference on those days, or the option of staying at home, unavailable to the participants in the shared activity. They are not rationally able to create plans to go somewhere else on their own or agree to do something elsewhere when they are supposed to be away with friends. The shared plan treated as exclusionary would filter out options that are incompatible with the shared plan. The intention to do something incompatible with the plan would mean that one is reasoning irrationally. This exclusionary treatment of plans, along with the other two norms of practical rationality, enables agents to reach their shared goal.

It is worth noting that treating plans as exclusionary does not mean that shared plans, and plans in general, are set in stone. Individuals are allowed to change their minds, if necessary, even when their intentions are part of a shared plan. To see how reconsideration works, refer to Chapter 3. Here I will add that, in cases of joint actions where agents do something with others, there will be expectations that those involved do their part to contribute to the joint project and coordinate their actions with other participants. Given these mutual expectations and predictions regarding the behaviour of others, abruptly dropping off a joint activity is bound to cause disruption and affect people's decisions and behaviour. Agents are free to

withdraw their participation in the joint endeavour. However, informing others that participation is about to be or has been withdrawn is advisable in joint activities, so that other participants can adjust their conduct and expectations accordingly. In some cases, changing one's mind and withdrawing from a joint endeavour without informing others might not be morally wrong. One of the cases in which it is wrong to drop out of a joint activity without warning other agents is the case of promises, where one promised another to do something together and the promise is broken without a stringent reason to do so, or without the promisee releasing the promisor from her obligation. Not all cases of joint activity, though, are based on promises or explicit binding agreements.²⁸

4.2 An explanatory model of shared practical reasoning

In this section, I put forward and analyse an explanatory model of shared practical reasoning. The model explains how agents are led to action when there is a shared plan in place that they are carrying out or will carry out with other agents. The section also presents a shared practical reasoning account that contains all the mental states and processes we discussed so far. In analysing individual actions and plans, I canvassed three explanatory models of practical reasoning, the belief-belief model, the belief-desire model and the belief-plan model. I also explained why I prefer the belief-belief model over the others. Here I rely on the latter to explain shared practical reasoning. If readers have different intuitions on what mental states prompt individuals to act on their plans, they can plug in a desire into this model or refer to the belief-plan model.

²⁸ Something to keep in mind is that dropping off a shared activity abruptly and without informing the other participants might lead to violations of trust. And some rich promissory obligations derive from letting others rely on you. So, an exploration of the connections between trust, relying on others and shared planning is necessary to understand when it is appropriate to drop off a shared plan without warning. I do not dive deeper into this exploration here.

The belief-belief model for shared practical reasoning can be formally rendered this way:

$$B(\text{sP } \varphi \text{ at } t) + B(\text{now is } t) \rightarrow \varphi\text{-ing}$$

This model states that what leads agents to act in the context of shared activities is a belief that there is a shared plan (sP) to do something at a time t coupled with a belief that the time has now come. These two mental states are individual beliefs in the mind of each participant in the shared activity. When there is a shared plan in place, participants will have individual beliefs that there is a shared plan to φ at t that involves them in some capacity that, coupled with a belief that t has come, prompts them to act by the content of the shared plan. $B(\text{sP } \varphi \text{ at } t)$ is a belief about what agents should rationally do when a shared plan involves them. Shared plans, like individual plans, constrain agents' thoughts and actions and are guided by norms of practical rationality. Norms such as consistency, means-end coherence and treating plans as exclusionary, require that an agent acting on a plan follows the norms themselves unless more stringent reasons to do otherwise arise. This will result in an agent making decisions and selecting options that are rational to make and select in light of the existing plan. A belief that there is a shared plan in place will influence what an agent decides or does next.

An important element to consider is acceptance of the shared plan by the participants and where acceptance fits in the belief-belief model for shared practical reasoning. Accepting a plan is fundamental to discharging it. Agents need to identify with the plan and recognise it as their own for them to act on it. Believing that there is a plan in place without accepting it, is equivalent to believing that someone, at some point, decided to plan something and that we are not involved in that plan in any capacity. The mere fact that a plan has been created does not prompt individuals to act if they do not see it as their own, namely, if they do not accept it. A plan to write a paper for the XYZ Conference, for instance, would not prompt me to take the relevant steps to write that paper if I did not acknowledge that the plan is in place for *me* to write a paper, that was created by me and that I accepted it.

In the case of individual practical reasoning, acceptance of plans is straightforward and tied to the belief that there is a plan in place. Upon planning something, the first stance an agent has toward the plan is acceptance of it. This is because plans are created to assist an agent in organising her conduct and carrying out complex goals. In doing so, plans help the agent to respond to reasons. It is rational, then, for the agent who created the plan to accept and follow it because responding to reason is rational, and plans help to respond to reasons. If one did not automatically accept a plan that one created,²⁹ one would be acting irrationally. Acceptance of plans, then, is always the first attitude an agent has toward her plan, where this acceptance can be revoked later if changes in circumstances or reasons require a change of plan.

Acceptance of shared plans works in the same way when group members, or simply individuals involved in a shared activity, equally contribute to the creation of the shared plan. When all agents have an equal say in how the plan is going to go and how its sub-plans are distributed among participants, acceptance of the shared plan is straightforward and tied to the belief that there is a shared plan, $B(sP \phi \text{ at } t)$. This is because, similarly to the case of individual plans, agents who have an equal say in the shared activity are co-creators of the shared plan and see it as their own. The fact that the plan is created by them makes their acceptance of it automatic. It is rational for them to automatically accept a plan they created, as that plan will help them to respond to reasons. Normally, agents who are equally involved in the creation of the shared plan have voluntarily joined a group that has shared goals to achieve, and that is why they are involved in the group's plans. This category of voluntarily joined, egalitarian shared activities encompasses familiar cases in the literature on shared agency, such as individuals painting houses together, going on walks together, and cooking together. It also

²⁹ Here, we need to distinguish between provisional plans and finalised plans. Provisional plans that agents might sketch out when they are surveying different options are not final and agents have not accepted them yet. However, when agents stop deliberating on what to do and settle on an intention or plan of action, that intention or plans are final. It is the mental state of settling the matter on what to do that causes the belief that there is a plan in place and, therefore, acceptance of the plan itself.

includes the case of the group of friends travelling together I used above. Groups of friends, clubs and associations are paradigmatic cases of voluntarily joined, egalitarian shared activities.

Acceptance of shared plans plays out differently in those cases where individuals do not equally participate in the creation of the shared plan. Some cases of agents acting together may not involve a stable group (one that persists through time) or identification with some shared goals, and individuals may agree to do something or join a certain group where shared plans have already been created. In those cases, they would not have a say in the shaping of plans for the group and they would follow a plan created by someone else. I argue that, in these cases, acceptance of shared plans is not automatic and is not tied to $B (sP \varphi \text{ at } t)$. Agents who join groups or tag along in shared activities where a shared plan is already in motion might reason this way:

$$B (sP \varphi \text{ at } t) + A (sP) + B (\text{now is } t) \rightarrow \varphi\text{-ing}$$

They believe there is a shared plan to φ at t ($B (sP \varphi \text{ at } t)$), they accept the shared plan ($A (sP)$) and, when the time comes, they act accordingly (they φ). This is how individuals are led to action when they participate in shared activities where they have not co-created the shared plan. Upon accepting the shared plan, agents will see it as their own, as something that affects and involves them in some capacity. Acceptance can be revoked later if circumstances change or if the agents have stringent reasons to reopen their deliberation and do otherwise. This non-automatic acceptance of shared plans might fit the case of protests, for instance, where we have a temporary group of agents gathered to make their voices heard. There might be a protest leader who has an agenda he is trying to push or a small sub-group of agents who initiated the gathering. However, as the protest grows and more people join, not all of them will have participated in the creation of a shared plan to protest on that day. Agents who join later might

infer that there is such a plan, perhaps they read the signs others are holding, and, upon reflection, accept it.

I now provide an account of shared practical reasoning that incorporates the main mental states involved in shared activities. The six-step account of egalitarian, shared practical reasoning unfolds as follows:

- 1) The group has shared goals to achieve.
- 2) Group members form end-related intentions toward the shared goal (e.g., I intend to do X with Y).
- 3) Group members form means-related intentions connected to their end-related intentions (I intend to do my part of X).
- 4) A shared plan is formed for the group to act on.
- 5) By default, group members must treat the shared plan (and its related reasons) as exclusionary.
- 6) 1-5 is common knowledge between participants

When participants in a shared activity have goals they want to collectively bring about, they can form end-related intentions toward their realisation. The content of these intentions is the outcome (goal) the agents' actions are jointly directed to. A division of labour usually goes on among participants in a shared activity, who can divide up the plan into smaller sub-plans that each participant can take care of. When this happens, each agent involved in the shared activity usually forms intentions concerning means to fill in the prior end-related intention (to fill in the prior plan). Once agents intend the ends and means a shared plan is in place for them to act on. By default, this shared plan and the reasons for action that are related to it must be treated as exclusionary. This is to avoid reconsidering the plan and sub-plans and to ensure the plan itself can successfully guide their conduct. If stringent reasons to do otherwise arise,

agents can reopen deliberation and stop treating that shared plan as exclusionary. All of the above is common knowledge between agents involved in a shared activity.

This is how multiple agents plan their conduct and act according to the shared plan to reach common goals. In the next section, I establish whether this model of egalitarian shared agency hinders the autonomy of the agents involved in shared activities.

4.3 Autonomy and the exclusionary nature of shared plans

In this section, I explore the implications of egalitarian, shared planning on individual autonomy. I focus on the norm of practical rationality according to which agents must treat shared plans as exclusionary. Does the fact that individuals treat shared plans (and the reasons they were created for) as exclusionary in contexts of shared agency make them less autonomous? Treating plans as exclusionary forces agents to stop deliberating on what to do and considerably restricts the options that are available for them to choose from. As I explained in Chapter 3, this does not constitute a problem for individual practical reasoning, which sees an agent independently devising her plans and acting alone. However, some constraints stem from shared plans, which I analyse in detail in this section. These constraints that group members treat as exclusionary are the product of different agents' practical reasoning. This means that, in doing things together, group members need to act in ways that are dictated by the shared plan, where they might not have decided on some or all aspects of the plan itself. We might think that the constraints that stem from shared plans impair agents' autonomy. I argue that this is not the case. Shared plans can help agents to respond to reasons and even enhance agents' reason-responsiveness capacity despite having constraints that are the products of multiple agents' reasoning.

Treating shared plans as exclusionary creates constraints that affect everyone who takes part in the shared activity. Different agents contribute to the creation of a shared plan's constraints, and those constraints influence the conduct of different agents. This means that when engaged in joint actions, an individual sometimes accepts and might act upon some reasons that were not entirely the product of her deliberation. Moreover, treating reasons as exclusionary prompts individuals to stop deliberating regarding what to do and restricts the options available to them. As a result of the constraints imposed by the exclusionary treatment of plans in shared activities, there is less room for deliberation for agents who act with others. An example might help us understand this better.

Let us consider HOLIDAY again. Suppose there is a group of three friends, X, Y and Z. Taking a trip together is the outcome they want to bring about by acting together (their shared goal). X, Y and Z meet up to choose a holiday destination and to finalise the details of their journey. They suggest two possible destinations, Edinburgh and Paris. Suppose they unanimously decide to put it up for a vote (following the majority rule). X and Y vote for Edinburgh while Z votes for Paris. The majority votes for Edinburgh, which then becomes the group's destination. Going to Edinburgh is now the group's shared plan and each agent intends to go to Edinburgh with the others. Once the plan is in place, X, Y and Z need to book their flights, find accommodation and create an itinerary of what they are going to do in Edinburgh. These are the sub-plans of their general plan. X, Y and Z decide that it would be less time-consuming and more efficient if each of them oversaw one sub-plan. So, they write down the three sub-plans (flights, accommodation, itinerary) on separate pieces of paper and put them in a bowl. X extracts the 'flights' piece of paper, Y gets the 'accommodation' one and Z the 'itinerary'. This prompts them to each intend their part of the shared activity: X intends to book the flights, Y intends to book their accommodation and Z intends to create an itinerary for their time in Edinburgh. These are their intentions concerning means that are connected to the end-

related intention that they each have to go to Edinburgh with the others. X, Y and Z then do their bit of the shared activity by each taking care of the sub-plan they picked up from the bowl. The group eventually brings about the outcome of going to Edinburgh together.

Let us analyse the situation from the point of view of Z. Before getting together with X and Y, Z has a reason to go to Paris and a reason to travel with X and Y. She has a strong preference for Paris as a travel destination and, if it was up to her, she would choose Paris as the group holiday destination. Before getting together with her friends to plan the group holiday, Edinburgh is not on Z's radar. Z's set of preferences before the shared plan exists is the following:

1. Going to Paris together
2. Going on holiday together
3. Going to Paris alone
4. Going on holiday alone

From the point of view of the other group members, both X and Y individually have a reason to go to Edinburgh and a reason to travel with the others before meeting up to finalise their holiday plans. Each of them has a strong preference for Edinburgh as a travel destination and, if it was up to them, they would choose Edinburgh as the group holiday destination. They also have a weaker preference for Paris. Let us assume that X and Y have the same sets of preferences. X's and Y's individual preferences before the shared plan exists are the following:

1. Going to Edinburgh together
2. Going to Paris together
3. Going on holiday together
4. Going on holiday alone

Once X, Y and Z meet up and form the shared plan to go on holiday to Edinburgh together, that plan gives them all a reason to go on holiday to Edinburgh with the others. By this, I mean that the plan reconnects them with the reason(s) it was created for. All of them had a reason to go on holiday together and Edinburgh is the chosen destination. The shared plan was created to respond to these reasons and, in following that plan, X, Y and Z are connected to the reasons that went into its creation. Since plans and their related reasons must be treated as exclusionary, they all need to abide by the plan and its constraints unless more stringent reasons to do otherwise arise. This means that X, Y and Z must each take care of a specific means to achieve the end to go on holiday together (e.g., booking flights, finding accommodation or creating an itinerary) and work towards bringing about the desired end (going to Edinburgh together).

X and Y each had a preference to go to Edinburgh with the others before the existence of the shared plan. However, Z did not have a preference or a reason to go to Edinburgh before. After the shared plan is created, Z has a reason to go to Edinburgh and must act on the constraints that the plan generated. She cannot make plans to do other things on the days of the trips and must plan an itinerary of things to do in Edinburgh with her friends. A shared plan produces constraints that are the product of different agents' practical reasoning. X, Y and Z pooled their preferences together and their decision-making processes produced the outcome of going on holiday to Edinburgh. In doing things together, group members need to act in ways that are dictated by the plan and they might also end up following other agents' reasons for action. In our example, Z goes to Edinburgh despite her original goal to visit Paris. This is because X and Y did not vote for Paris. Z has indeed a reason to accept the result of the majoritarian voting procedure, given that, as we assumed, she voluntarily subjected herself to it. However, it is also true that Z has now a reason to go to Edinburgh that she has only as a result of the shared plan. As an individual, Z would not deliberate and decide to visit Edinburgh.

Is Z less autonomous in following a reason that stemmed from the shared plan, if she did not have that reason before?

On the face of it, individual autonomy seems jeopardised in cases such as the one above where an individual is responding to someone else's preferences and does not act on her preferences. Being autonomous means a) deciding for yourself what to do, and b) responding to reasons. When agents participate in shared activities, they often do not make (all) decisions related to the joint activity and are influenced by what other participants choose as options for the shared plan. In the holiday case, each friend takes care of one aspect of the trip so Z, for instance, does not directly make decisions regarding flights or the hotel she will be staying at. Moreover, she is influenced by what X and Y choose as options. X and Y propose a different holiday destination which gets chosen over Z's preferred holiday destination. Also, agents involved in shared activities seem to respond to reasons that are not theirs, since they end up responding to some of the other group members' reasons. Z, for instance, responds to X's and Y's reason to go to Edinburgh when she acts on the shared plan to go to Edinburgh with the others. Z did not have that reason before.

Shared plans do not impair individual autonomy, for they help agents respond to reasons. I expand on this point now. Later I address whether agents should always be the authors of their decisions to be autonomous.

There are two ways in which we can show that shared plans do not endanger the agents' autonomy in cases of acting together. Firstly, the reasons of the group can be or become the reasons of the group members. Individuals have a reason to join groups or to be group members. This reason is that they share goals with the agents that form a group.³⁰ Those goals are best achieved via cooperation with others. That cooperation may be shaped and facilitated

³⁰ Agents may have individual goals that coincide with the group's goals. But they may also want to join a group for its own sake and later the group's goals become the agent's goals as well.

by the creation of a shared plan and the relative sub-plans. In this case, the reasons that stem from the shared plan are also the group members' reasons because they are part of a shared activity that is motivated by goals that all the members share. If that is the case, group members are still autonomous in similar contexts. X, Y and Z have a reason to be part of the friend group, and that reason, we assume, is related to their mutual friendship and affection. They all share goals with the others, such as spending time together as friends or, specifically, taking a trip together. The goal of taking a trip together is best fulfilled via cooperation, cooperation that is aided by forming a shared plan. The reasons or constraints that stem from the plan are related to the goals and reasons they individually have already (e.g., spending time with their friends). So, they remain autonomous in following the shared plan.

This argument, however, only accommodates those cases in which individuals identify with some goals they share with others. Groups of friends, clubs and associations are all examples of groups where people identify or come to identify with other agent's goals or endeavours. Sometimes, group membership produces identification of the individual with the group and this can be a condition for the group's reasons to become its members' reasons. It is not plausible to say, however, that membership and identification with the group are at the basis of all or most shared activities.

Some cases of agents acting together may not involve a stable group (one that persists through time) or identification with some shared goals. My neighbour N, for instance, asks me for help in painting her house. She has already started painting the front and wants me to do the back. N has a goal, 'properly painting the house', which I do not share (I am indifferent as to whether her house is painted). N also chose paint that is notoriously toxic for the environment, and I do not approve of that. However, I decided to help her for a purely instrumental reason. N has a pool and I want to use it in return for my help. It is true that, by helping her, I voluntarily joined this one-off group. However, I neither identify with it nor share

its goals or values because my association with N is purely instrumental. Despite my transactional association with N, I am now part of a shared plan to paint N's house. Some constraints (using the toxic paint she bought) and choice points (N started painting the front of the house and now I have no choice but to do the back) stem from the shared plan and exclude my deliberation. Am I less autonomous in painting the house with N?

Shared plans might help individuals to respond to reasons even when they do not share goals with the group and when they are acting for different reasons than other group members. An agent may have a reason to do something. In instances of shared activities, coordinating and organising our conduct with others may help us to do what we have a reason to do. We do better by organising and coordinating our conduct via plans. Therefore, shared plans may help us to better respond to reasons for action we already have. The exclusionary element of shared plans does not prevent an agent from exercising her autonomy if that shared plan helps the agent to better conform to her reasons. In the previous example, I have a reason to use N's pool. Painting the house with N is a shared plan that enables me to respond to my reason. Even though I do not share a goal with N, I do not care about doing a good job and I am subjected to constraints that follow from the shared plan, I am not less autonomous in acting with N. The shared plan's filtered options help me to respond to my reason to use N's pool.

So, when we treat shared plans as exclusionary, we stop deliberating about courses of action and that reduces the options that are available for us to act on. This does not seem to threaten personal autonomy. Being autonomous, among other things, entails responding to reasons. Shared plans, by organising and coordinating agents' conduct in the context of group action, allow agents to respond to reasons that they share with the group and personal reasons they have as individuals.

What about deciding for oneself? Autonomy has a strong decision-making component, for it requires individuals to decide for themselves how to act. As I mentioned in Chapter 3, though, deciding for ourselves does not mean that we have to continuously deliberate and make decisions at every turn, or that we need to be the ones deciding everything. In the individual case where we act by ourselves, we can rely on plans made by past time-slices of ourselves without reopening deliberation. We can also follow someone else's advice and allow that advice to guide our decisions and so forth. That enables us to save time in the decision-making process and respond to reasons. Similarly, in acting with others, agents do not need to decide every aspect of a shared plan to be autonomous. They can take care of certain sub-plans while other agents take care of others, join a group and act on its ready-made plans at any point and rely on other group members' judgement to make personal decisions. When agents decide to curate an aspect of a shared plan, start following a plan that was made by someone else or rely on others to make decisions, they are going to have mental states and reasons that stem from the decisions of others. For instance, they are going to have beliefs that they have a sub-plan to take care of or a belief that they are now following a ready-made plan. These beliefs will most likely prompt them to form relevant intentions, such as intentions to act on the sub-plan, or intentions to do what the shared plan they joined requires them to do. The sub-plans or the shared plan they joined will also provide reasons. All these mental states – beliefs, intentions, reasons – will enter individuals' deliberative processes and interact with other mental states individuals might have. Agents will, as a result, make decisions based on those mental states and adapt their conduct accordingly. In a way, agents decide for themselves what to do even when they are not the creators of a plan or the sole decision-maker of a shared project.

All of this enables individuals to respond to reasons while fully living and operating in societies or group contexts, where agents must cooperate to have relevant forms of sociality.

Personal autonomy, then, is fully maintained in contexts of shared agency where plan-related reasons are treated as exclusionary.

I will now make another point when it comes to the relationship between personal autonomy and cases of shared agency. If we accept that separate time-slices of the same person can act on plans and bring about complex goals without hindering personal autonomy, we need to accept that agents can create shared plans and act accordingly without negative consequences for their autonomy. Separate time-slices of the same person are, in a way, similar to different agents. They are somewhat separate entities that are connected through relations of consistency and coherence between intentions and plan-related mental states. Past time-slices can make plans that sometimes stretch far away in the future and influence the thoughts and actions of several future time-slices. When future time-slices enact plans made in the past, their deliberative processes are facilitated. They do not have to create a plan from scratch and can rely on the organisational work done by past time-slices to reach their complex goals. Plans created by past time-slices, then, help future time-slices by making their deliberation and reason-responsiveness easier. Plans made in the past, as we know from Chapter 3, can help individuals respond to reasons in the future. The same mechanism is at play when people act together with others. Shared plans make deliberation easier for individuals, who do not have to take care of every organisational aspect of a project or plan by themselves. Shared plans also help their participants respond to reasons, as argued above. If we accept that different time-slices of the same person can ‘work together’ toward the realisation of a personal project in the future, we need to accept that different agents can do the same to bring about shared goals and discharge shared plans. Neither of those two cases threatens personal autonomy.

One might argue that there is a substantive difference between ‘acting with’ past time-slices of oneself and acting with other agents and that this difference matters for autonomy. When I act on my past time-slices’ reasons, intentions, and plans, I am acting on my own

reasons, intentions and plans, for past time-slices of myself are connected with my present time-slice and future time-slices to form a somewhat unitary self. In acting with and relying on past time-slices of me, I am only responding to and relying on myself, with no other agent involved. When I act with others, I need to account for other agent's reasons, intentions and our shared plans and I rely on other individuals. These two cases could not be more different, or so it seems.

To understand my analogy better, let us draw a parallel between reliance on past time-slices and reliance on other agents. Reliance on oneself is fundamental because we cannot make sense of autonomy without it. As rational agents we often want to pursue complex goals and, to do that, we need to rely on past time-slices' plans and intentions. If we lived on a moment-by-moment basis, where time-slices of ourselves are separated from one another, we would not be able to do anything that requires extensive planning in the past for future actions. But we are capable of such planning, which means that we continuously refer to and communicate with past time-slices of ourselves. Defining autonomy without this reliance on past time-slices of oneself would not make sense of how we act and think of ourselves as autonomous agents.

A conception of autonomy that excludes reliance on other agents (the one that Wolff's anarchism seems to point to) is also unattractive. It confuses self-determination – autonomy as making your own decisions and responding to reasons – with complete self-reliance – autonomy as relying exclusively on yourself for every action or decision. We live in highly interconnected societies where, even when we act by ourselves, we depend on others to do certain things. For instance, as an autonomous agent, I decide to send my mother a pottery mug as a gift. The success of my plan is dependent on the manufacturer of the mug sending the stock to the store, the store's employees regularly stocking shelves, other customers' not getting to the last mug before I do and the post office's employee putting the right label on the parcel addressed to my mother. Cases like this, where individual decisions and actions are dependent

on the decisions and actions of others, are very frequent in societies. Relying on other agents to do their part is a big component of autonomy because, unless I made and delivered the mug myself, I could not gift it without relying on the manufacturer, the store's employees, other customers and the post offices' employees. Most, if not all, individual actions require reliance on others and only a conception of autonomy as complete self-reliance would ignore this. Being autonomous, however, does not mean that we do everything by ourselves without any external influence. It simply means making decisions and being responsive to reasons and if other agent's actions help us in this process that does not make us less autonomous.

The argument above goes from the reasonableness of relying on past time-slices to the reasonableness of relying on other people (in acting together with them or not). But to illuminate the notion of autonomy is also useful to invoke the argument that goes in the other direction. The kind of reasons Wolff has for rejecting reliance on others also seem to cast doubt on reliance on past time-slices of oneself. The main problem Wolff has with relying on others is that relying on other people is a kind of irresponsibility. By doing that, one abandons (full) responsibility for one's own decisions. Wolff is attracted to the idea that, when one is deciding what to do, one must decide for oneself. To do what one is told to do is, for Wolff, to abdicate responsibility for deciding for oneself. The same point, however, can be made about reliance on past time-slices of oneself. When I follow a plan made by a past time-slice, I might be abdicating responsibility for deciding what to do *now*. There is no guarantee that in this moment I would come to the same decision as my past self did. This argument, which goes from the (un)reasonableness of relying to other people to the (un)reasonableness of relying on past time-slices of oneself, exposes the absurdity of interpreting autonomy as Wolff does. An understanding of autonomy that prohibits reliance on my past self is absurd, for it would be impossible for us to bring about any complex goal that requires that we perform a series of actions overtime.

The fact that our lives are extended overtime and the fact that we live social lives should both be accommodated in a theory of autonomy. Excluding reliance on other agents as part of autonomy is as bad as excluding reliance on past time-slices. Most autonomous actions could not be completed without them. Wolff's (1970) conception of autonomy does not allow him to accommodate cases of acting together without individuals losing their autonomy. This is because he defines autonomy as complete self-reliance. I showed, though, that autonomy more neutrally understood is compatible with joint action in non-hierarchical settings.

Lastly, it is important to highlight that shared plans enhance agents' reason-responsiveness capacity. In the individual case, when we consider one agent acting over time, future time-slices of the same self have access to all the reasons of past time-slices via plans. When a time-slice creates a plan that stretches over time, that plan, which affects several future time-slices, acts as a bridge to the original reason(s) that prompted the past time-slice to plan. As a result, the reason-responsiveness capacity of future time-slices is enhanced. Future time-slices can respond to reasons past time-slices considered before and free up some of their cognitive space. This way they will bring about complex goals they would otherwise not be capable of bringing about without 'working' with other time-slices. Similarly, in the context of shared agency, group members have access to all the reasons of the group via shared plans. Shared plans function as vessels that 'contain' the reasons they were created for. Group members who co-created a shared plan, as well as agents who subsequently join the group and act on its plans without creating them, have access to all the reasons the shared plan was created for. They will not have to deliberate by themselves to initiate and complete complex projects, which saves them time and cognitive effort. With the help of shared plans, agents can fulfil shared goals and do things that they would not be able to do without collaborating with other agents.

So, shared plans not only help agents to respond to reasons, but they also augment individual cognitive capacities by giving agents access to more reasons. These plans provide agents with the tools to deliberate on what to do collectively in a way that saves them time, effort and resources. Individuals do not have to do all the cognitive work that goes into planning by themselves: they can divide up the labour among the group, respond to reason without complex and time-consuming deliberation and bring about goals they could not achieve by themselves. This is how shared plans protect and enhance personal autonomy.³¹

4.4 Summary

In this chapter, I analysed shared practical reasoning, that is, how agents come together to create shared plans that allow them to reach their goals. In section 4.1, I canvassed the main mental states agents have when they act with others and the norms of practical rationality that govern shared activities. In section 4.2, I provided an account of shared practical reasoning that explains how agents come to form a shared plan and act on it to reach common goals. Finally, in section 4.3, I investigated whether shared plans negatively impact individual autonomy. In doing so, I focused on the norm of practical rationality according to which shared plans and plan-related reasons must be treated as exclusionary. I argued that, when treated as exclusionary, shared plans generate constraints that are the product of different agents' practical reasoning. In following the shared plan, agents might end up acting on someone else's reasons

³¹ Here I'm arguing in favour of a view where agents acting together do not have to do all the cognitive work that goes into planning a shared activity by themselves as individuals. They can instead divide up that cognitive labour amongst the group members and respond to reasons without re-deliberating on what to do at every turn. In these instances of joint actions that rely on a division of labour among different agents to create and execute shared plans, agents might respond to their own reasons or, even, the reasons of other group members. Agents' autonomy is protected, as explained in this section, since they still make decisions and decide for themselves when they incorporate other agents' reasons in their own practical reasoning. This ability agents have to respond to their reasons and other agents' reasons hinges on a mechanism which I call 'transmission of reasons'. I explain transmission of reasons and compare it to other accounts in the action theory literature in Chapter 5, section 5.1.1. There, I also offer a defence of transmission of reasons against the classic, Davidsonian account of reasons.

for action. Despite this, I argued that shared plans do not impair individual autonomy for two main reasons.

Firstly, the reasons of the group can be or become the reasons of its members. When this identification with the group's reasons occurs, the reasons that stem from the shared plan are or become the group members' reasons because they are part of a shared activity that is motivated by goals that all the members share or come to share. In following plans in these cases, individuals do not lose their autonomy because they respond to reasons they already have or come to have when they identify with the group's endeavours. However, not all cases of joint activities display such convergence of reasons between group members or identification with the group. However, shared plans might help agents respond to reasons even when they do not identify with the group or are acting for different reasons than other group members. In some cases, coordinating and organising our conduct with other agents might help us to do what we have a reason to do. Since we best organise our conduct via plans, shared plans may help us to better respond to our reasons even when we do not identify with the group or when we are acting for different reasons than other agents. This is how individuals maintain their autonomy when following shared plans created in the context of egalitarian shared agency.

Chapter 5 - Hierarchical Shared Agency & Hierarchical Groups:

How Authorities and their Subjects Do Things Together

5.0 Introduction

In this chapter, I analyse hierarchical shared agency, that is, those instances of acting together that involve asymmetries of power in a group, where one or some group members have and exercise authority over other group members. Hierarchical shared agency occurs when there is a transmission of reasons from one or more agents to others. In this chapter, I explain what reason transmission, compare this account to other approaches in the literature and defend it against criticism.

The chapter has three main aims. Firstly, it aims to analyse hierarchical shared agency and how reasons transmitted from agent to agent can motivate them to act and explain what they do.

Secondly, the chapter offers an analysis of nation-states as structured groups of agents acting together. It delves deeper into the status roles and functions group members have in nation-states. Lastly, it investigates whether personal autonomy is compatible with hierarchical shared agency. This constitutes the third stage of the three-stage model of practical reasoning to solve Wolff's autonomy-authority problem.

The chapter is divided into four main sections. Section 5.1 is dedicated to the analysis of transmission of reasons, hierarchical shared agency and group hierarchy. Section 5.1.1 explains what reasons transmission is, how it occurs and how it differs from other approaches in the literature. It also offers a defence of transmission of reasons against objections. Section 5.1.2 defines hierarchical shared agency, explaining the conditions that need to be satisfied for it to arise. Section 5.1.3 addresses the internal structure of nation-states as hierarchical groups, with

emphasis on the roles that agents occupy them and the notion of status function (Ludwig, 2017). Section 5.2 deals with how political authorities in nation-states plan for their subjects. It presents a six-step account of hierarchical shared agency that explains how political authorities create plans and how their subjects incorporate them into their practical reasoning. Section 5.3 goes back to Wolff's autonomy-authority problem to investigate whether acting autonomously is compatible with abiding by the authority's plans. I then offer concluding remarks in section 5.4.

5.1 Hierarchical Shared Agency

In this section, I analyse transmission of reasons and hierarchical shared agency. Section 5.1.1 presents and defends my account of reason transmission. Section 5.1.2 defines hierarchical shared agency and clarifies the conditions that need to be satisfied for it to arise. I distinguish it from group hierarchy, which is the internal structure of authority relations some groups, such as nation-states, have. Section 5.1.3 addresses the internal structure of nation-states as hierarchical groups.

5.1.1 Transmission of Reasons

I argue later that for a hierarchical, shared activity to occur, there needs to be a transmission of reasons from one person to another.³² This means that someone needs to make other people act for her reasons. Here, I canvass my account of reasons transmission from agent to agent. In section 5.1.2, I explain hierarchical shared agency.

The main idea behind transmission of reasons is that intentions and plans are means of transmitting reasons from agent to agent. Plans, and the intentions that form them, are

³² I am indebted to Abraham Roth (2017) for the idea that reasons can be transmitted from one agent to another. He calls this 'entitlement to another person's reasons for action' and explores this idea in the context of egalitarian, shared agency when people act together without authority relations amongst them. Here I explain how this might happen in hierarchical group settings and, among other things, compare my account to Roth's.

‘repositories’ of reasons, metaphorical vessels that store the reasons they were created for. Agents create plans of action as a result of having specific reasons that motivate them to act. Those plans retain the reasons they were created for until agents complete the relevant tasks.

I started analysing plans as vessels that store reasons for action in Chapter 3 when I discussed individual practical reasoning. An agent might have reasons to do something and can create a plan to help her respond to those reasons. That plan stores the reasons for action and the agent can respond to them at any point in time by fulfilling that plan. In certain contexts, including group contexts, some actions might need to be carried out by multiple agents. When agents act together, shared agency occurs. However, if someone (or several people) transmit her reasons for acting in a certain way to others, either by expressing her intention that others do something or by creating a (detailed) plan, and other people act on those reasons (by acting on the intention or following the plan), we have an instance of hierarchical shared agency. In these cases, agents make people act on reasons they have, and they will in most cases also coordinate their conduct. The agents who transmit reasons to others via plans or intentions might take charge of a shared activity and have the power to influence how others are going to act. The others will, then, might be in a subordinate position when they do as they are told. So, what constitutes hierarchical, shared activity is a transmission of reasons from one or more agents to other agents. (p.123).

By transmitting its reasons to others, authority can guide them the way their own reasons would. The point of this reason-transmitting mechanism is for us to treat the authority’s reasons as motivating reasons for action, in such a way that they feature prominently in our practical reasoning by guiding and directing our conduct. The idea that we treat – or should treat – the authority’s reasons as our motivating reasons for action comes from Raz. I explain what this treatment of authoritative reasons amounts to if we are to take Raz seriously on this point.

Raz (1990) argues that authorities (i.e., practical authorities) can make others act on their reasons. For Raz, practical authorities provide others with reasons to behave the way they are telling them to behave, and these reasons need to be taken as action-guiding by others (Ehrenberg, 2011). This is showcased in Raz's pre-emptive thesis (Raz, 1986), which is a thesis about how subjects should treat authoritative directives (Ehrenberg, 2011, p. 887). For Raz, the authority's reasons pre-empt the subject's reasons for action in the sense that, when the authority requires the performance of an action, that request is a reason for doing that action that excludes and replaces other reasons the subject might have for not so acting (Raz, 1986, p. 46). The subjects should, then, allow authoritative directives to be rules that guide their actions, and this requires treating those rules as motivating reasons (Ehrenberg, 2011, p. 887).

Motivating reasons are reasons for which we do something, reasons that, in our eyes, count in favour of us acting in certain ways (Alvarez, 2017). In this light, the pre-emptive thesis is truly about how the authority's reasons become our motivating reasons for action. If we are to take Raz seriously on this, we need an account of how we can practically incorporate the authority's reasons into our practical reasoning.

However, if we are to follow Raz on how to treat authoritative directives, we cannot forget that, on the service conception, his pre-emptive thesis is connected to what he calls the dependence thesis. According to the dependence thesis, authoritative directives need to be based on pre-existing reasons which apply to the subjects independently of the existence of the directives (Raz, 1986, p. 47). We might think that when the authority's directives are based on the subjects' reasons, it should be easy for the subjects to take the reasons behind the authoritative directives as their motivating reasons for action. After all, the directives are based on reasons the subjects already have.

However, there is a difference between the reasons the subjects have and the reasons that apply to the subjects. Having a reason is traditionally taken to mean that an individual represents that reason in her mind in a propositional form and, in some cases, sees that reason as favouring a course of action.³³ In paradigmatic cases, when someone represents a reason in her mind and takes it to favour an action, that reason is a motivating reason that is available for her to act on. However, some reasons can *apply* to an agent without the agent representing them in her mind or being motivated by them (Alvarez, 2017). This might be the case for moral reasons, as morality might require that we do something regardless of us representing that moral reason and that reason motivating us. So, we might be motivated by reasons that apply to us (and represent them in our mind). But, in some cases, we do not represent and are not motivated by a reason that, nevertheless, applies to us. When there is a discrepancy between reasons that apply to us and reasons that we represent (or motivate us), we need to explain how the reasons that apply to us can come to motivate us – and prompt us to act - if we do not represent them in our mind. In the authority case, when the reasons behind authoritative directives already apply to us without us representing them in our minds, we need to see exactly how those reasons can come to motivate us and make us act on them. So, even though authoritative directives are based on the subjects' pre-existing reasons, we still need to explain how the subjects come to treat those reasons as motivating reasons for action despite not representing them in their minds. According to the pre-emption thesis, this representation happens with the mediation of the authority. It is through the authority's reasons that we reconnect with reasons that already apply to us.

³³ This is, for instance, Davidson's view (1963). I am sceptical that this is the only way in which we can have reasons or in which reasons apply to us. Transmission of reasons is an alternative to this traditional way of thinking, an alternative according to which we do not always need to represent a reason in our minds in propositional form for that reason to motivate us to act.

The challenge that we face in explaining how the subjects can treat someone else's reasons as their motivating reasons for action stems from the belief that reasons need to be possessed by the agent who is acting on them to explain their intentional actions (Anscombe 1963; Davidson 1980). By possessing, it is generally meant that reasons need to be represented as propositions in an agent's mind. When we act on reasons, traditional accounts say, we represent related goals in our minds and create plans that would allow us to reach them (Roth, 2017). This traditional theory can explain how we act on our reasons, as we represent the reason in our minds and that representation, in the form of a proposition, can guide our thoughts and actions in a relevant way. However, this traditional approach to reasons does not explain how the authority's reasons can guide us if we do not directly represent them in our minds.

I offer a model of how the authority's reasons can become our motivating reasons even when we do not represent them in our minds. I argue that the authority's reasons can be directly or indirectly accessible to the subjects. In both cases, the authority communicates its intentions (or plans) to the subjects and, through that communication, the reasons behind its directives (plans) are transmitted to the subjects.

The authority's reasons are directly accessible to the subjects when the authority openly communicates the reasons behind its plans to the subject. For instance, suppose that the spread of the COVID-19 virus can be decreased by 50% with a national lockdown. Call this the fact that *p*. The authority knows that *p*. As a result of knowing that *p*, the authority has a reason, *R*, to reduce the spread of the COVID-19 virus with a national lockdown, presumably connected to a general reason to protect its subjects' health. The authority creates a plan that instructs its citizens to stay at home. This plan gives the subjects a reason to stay at home. The authority imposes the plan on its subjects and also communicates that *R* is the reason behind the plan and that *R* = the fact that *p*. The subjects act on the plan by staying at home.

When the subjects follow the plan to stay at home they act on R (the reason to reduce the spread of the COVID-19 virus with a national lockdown), which is the authority's reason for creating the plan. The subjects have access to and can act on R based on testimony. When the authority communicates the plan to them, they come to believe, based on that communication, that there is a reason R which prompted the authority to create the plan. Given that the authority explains all the relevant facts, the subjects also know that R. When someone asks them why they are staying at home, they can answer that they are doing so not simply because there is a plan in place that demands that they do, but because of p. They have not personally investigated the fact that p, but they have come to know it through the authority's testimony. The plan to stay at home gives the subject a reason to stay at home. In this case, the subjects know the authority's original reason, and that R = the fact that p. With the authority's testimony, the authority's reason is directly transmitted to the subjects and becomes a motivating reason for them to act on. The authority's reason will, then, favour the subject's actions when they follow the authoritative plan. Through this direct communication, the subjects can represent the authority's reason in their minds.

Knowledge or belief about the authority's reasons based on testimony requires clear communication of those reasons on the part of the authority. Not all instances of an authority planning for its subjects, though, are backed up by that kind of communication. Sometimes the authority issues directives without providing a clear explanation of the considerations that went into formulating those directives. In this case, the authority's reasons can also be indirectly accessible to the subjects. When this sort of explicit communication on the part of the authority does not occur, the subjects can access the authority's reasons indirectly via the authority's intentions and plans. As I explained before, plans, and the intentions that form them, are 'repositories' of reasons. They are created as a result of an agent having specific reasons for action and they retain those reasons for the time they are in place. An agent can act on those

intentions or plans in the present or the future and, when that happens, she acts on the reasons ‘contained’ in the intention or plan. Similarly, an authority might have reasons that agents do something, can create a plan to help them respond to those reasons and communicate that plan to the subjects. That authority’s plan stores the reasons for action the authority had to create it, and the subjects indirectly respond to those reasons by fulfilling that plan (i.e., by obeying the authoritative directive).

To put it differently, the authority knows the fact that *p* and has a reason as a result of knowing *p*. The authority creates a plan to respond to that reason and the reason is, then, ‘stored’ in the plan. The authority communicates the plan to the subjects and, when they act on the plan, they indirectly act on the authority’s reason that went into the creation of the plan. Here, the authority’s plans function as motivating reasons for the subjects. When the subjects are aware of the authoritative plans, they represent the plans’ contents in their minds (e.g., ‘pay your taxes by April’). Given that the plans store the reasons they were created for, by acting on the plan, the subjects also act on the reasons behind the creation of the plans, albeit indirectly. The subjects come to (directly) know the reasons behind the plans when the authority conveys that information.³⁴ So, reasons for action are transmitted from agent to agent either directly via testimony or indirectly via the sharing of intentions or plans.

Arguing that reasons can be transmitted from agent to agent, or from an authority to its subjects, goes against how philosophers normally think about acting for a reason. However, it has some precedents in the literature. Abraham Roth (2014; 2017; 2020) has investigated the possibility of our behaviour being explained by someone else’s reasons in the context of egalitarian, shared agency. He argues that, in some cases, we are entitled to another agent’s reasons for action and, in virtue of that entitlement, we can act for reasons that we do not

³⁴ Public reasoning can also help the subjects access the authority’s reasons. I analyse public reasoning on the part of political authorities in Chapter 6.

possess, that is, reasons that we do not represent in our minds (Roth 2017, p.76). For agents to be entitled to each other's reasons for action, according to Roth, they need to act together to reach some common goals. When they act together towards common goals, and one instructs someone else to do something, the person receiving the instructions is acting on the reasons behind the instructions. That reason is possessed by the other agent and transmitted via the communication of an intention. Roth presents a case where we are going to a conference downtown by car, where I drive and you navigate (Roth 2017, p. 82). When you ask me to turn left, I turn left. The reason why I am turning left is, for Roth, not simply because you asked me to. I am turning left in order to follow the route to get downtown even though you know what the exact route is and I do not. Your intention that I turn left, which you communicate to me, contains your reason for turning left (e.g. that is the fastest way to get us to street X that will take us downtown) and, in virtue of us acting together on the common goal of getting downtown, I am entitled to your reason. Despite not representing that reason, I can act on it, and that reason can explain why I am turning left.

The fact that I can be entitled to the reasons of another and that those reasons explain my actions, can be also seen in cases of diachronic, intrapersonal actions. When an agent is engaged in individual planning agency across time, the agent initially represents a reason in her mind and, as a result of that reason, creates a plan. At some point during action and decision-making, the reason that prompted the agent to create the plan does not need to be represented for the agent to be able to act on it (Roth 2017, pp. 82-83). Once the deliberation has been made, the agent has weighed that reason against others and an intention (and plan) has been formed as a result, the reason takes a backseat in the agent's mind and the agent can act on it by following the related plan. A future time slice of the same agent is entitled to her past time slices' reasons when it acts as the past time slices prescribed. The reasons of the past time slices explain why the future time slice acted that way even though the future time slice does not

actively represent the past time slices' reasons at the time of action. Roth believes that this extends to the interpersonal case and that agents who act together towards a common goal are entitled to each other's reasons.

I am sympathetic to Roth's bold and innovative account of shared practical reasoning. His account differs from my account of reason transmission on two grounds. Firstly, Roth appeals to explanatory reasons for action, whilst I appeal to motivating reasons. According to Roth, what sometimes explains my action is the reason of another which I do not represent in my mind at the time of action, but I am entitled to under certain conditions. According to my account of reason transmission, the reason of another can be metaphorically transmitted to me via that person's intentions and plans. When this occurs, I argue that such reason motivates me to act as if it was my reason and figures in the explanation of my intentional actions even though I do not represent it in my mind when I act. This concept of transmission of reasons is inspired by Roth's early work on this topic. In the early stages of setting up his theory, Roth argued that agents can act on the intentions of another when they work together towards bringing about a common goal (Roth, 2004). Agents can, so to speak, directly execute the intentions of another when those intentions are communicated to them. For instance, when you say to me "Turn left" and I execute that directive, your intention to turn left is enacted by me when we drive to the conference together. Later, Roth appeals to reasons rather than intentions, arguing that someone else's reasons can explain my actions in acting together. With my reason transmission account, I essentially merge Roth's references to intentions and reasons, and this is the second difference between our theories. I argue that a causal connection between reasons and plans is formed when we create a plan of action as a result of a reason we have. When this happens, the plan exists exclusively because we have a certain reason and would not have existed otherwise. This causal connection between reasons and plans enables plans to 'store' the reasons they were created for. We, then, act on those reasons when we discharge the plan,

either right now or in the future. So, the main difference between Roth's account and mine is that Roth is interested in what explains my actions, while I am interested in figuring out what motivates me to act. What explains my action (for Roth) and what motivates me to act (my account) can be, in some cases, the reasons of another. Also, for Roth, I am entitled to your reason, which means I can point to it to explain my behaviour. I argue instead that your reason is transmitted to me via your intentions and plans in acting together.

These nuances differentiate transmission of reasons from Roth's theory of entitlement to another's reasons. The two accounts, though, are very similar. Roth focuses on entitlement to someone else's reasons in the context of egalitarian shared agency. I believe we can generalise his approach to incorporate cases of hierarchical shared agency, where agents act together under the authority's direction. Roth states that there needs to be a "causal and historical story to be told" for entitlement to someone's reasons to arise (Roth 2017, p. 83). We can reconstruct this causal and historical story to extract the ingredients to entitlement. When two or more agents act together (without any authority relationship) towards a common goal, they can create a plan of action. In the case of the two agents driving downtown, if one of them is tasked with figuring out the route, that person can create a route-related sub-plan as a means to the end of getting downtown. That agent would represent the route-based reasons in her mind and use them to create the relevant sub-plan. When the agents are in the car together, enacting their plan to get downtown, the driver is entitled to the other agent's route-based reasons because these reasons – and their related sub-plan – serve the common goal of getting downtown and the agents agreed to split certain tasks (one agent is in charge of driving, while the other is in charge of finding the best route).

So, the ingredients of entitlement to another agent's reasons are acting together, the presence of a common goal, a division of labour amongst the agents and a past deliberation on the part of one or some agents as a means to reach the common goal. These ingredients are

present in cases of hierarchical shared agency between political authority and its subjects, making it possible to generalise the notion of entitlement to include hierarchical shared agency.

Political authority and the agents under its jurisdiction can act together toward common goals. These common goals tend to be solving coordination problems and prisoner's dilemmas, reaching and maintaining social justice and so forth. Furthermore, hierarchical shared agency is based on a division of labour amongst the agents involved. The agents in authority create shared plans to reach said goals and the subjects are expected to follow the authority's plans and be sensitive to them when making personal plans.

When the authority and its subjects are involved in shared agency, they can create plans to reach common goals. In large-scale groups, division of labour is paramount to acting together. The authority is tasked, via delegation on the part of the subjects, with trying to solve social issues, and the subjects are tasked with following the authority. To reach common goals, the authority creates plans as a means to reach those goals. The authority represents reasons for action in its mind and uses them to create plans for the subjects. When the subjects abide by the authority's plans, they are entitled to the authority's reasons because those reasons – and the plans that resulted from them – serve the common goals and the agents involved (authority and subjects) agreed to that division of labour.³⁵

If the authority explicitly communicates the reasons that went into the creation of its plans to the subjects – as in the case of testimony discussed earlier - the subjects would have direct access to the authority's reasons via entitlement. When this does not occur, the subjects do not know what the authority's reasons are but can access them indirectly via the authority's intentions and plans. Plans, and the intentions that form them, are 'repositories' of reasons. They are created as a result of having specific reasons for action and they retain those reasons

³⁵ I will examine the extent to which this agreement is voluntary in discussing delegation in Chapter 7.

for the time they are in place. An agent can act on those intentions or plans in the present or the future and, when that happens, she acts on the reasons ‘contained’ in the intention or plan. Similarly, an authority might have reasons that agents do something, can create a plan to help them respond to those reasons and communicate it to the subjects. That authority’s plan stores the reasons for action the authority had, and the subjects respond to those reasons by fulfilling that plan (i.e., by obeying the authoritative directive). Here is how we can generalise Roth’s account to incorporate cases of hierarchical shared agency. So, my transmission of reasons account differs from Roth’s on the two grounds described above and includes cases of acting together with authority relations in the group.

Other accounts in the action theory literature explore similar issues with more caution. An alternative solution was developed by Saba Bazargan-Forward (2022). Bazargan-Forward is interested in how to attribute individual accountability for wrongful actions that are undertaken cooperatively, with other agents (2022). He contends that sometimes, to morally evaluate an agent’s conduct, we need to refer to someone else’s reasons for action. The other agent’s reasons determine the purpose of our conduct and, if we are doing something wrong, the other agent is accountable for our wrongdoings. He illustrates these points with *Olympic Sabotage* (Bazargan-Forward, 2022, p. 35). In *Olympic Sabotage*, Contender and Victim are both eyeing first place in the upcoming marathon. To reach this goal, Contender delegates to Goon the task of harming Victim just enough to get her out of the competition. Goon promises to oblige, and Contender also accepts this promise (Bazargan-Forward, 2022, p. 35). In this scenario, Bazargan-Forward contends, we can hold Contender’s accountable for Goon’s immoral conduct because Contender furnished Goon with the purpose of attacking Victim. Goon is of course also accountable for their wrongdoings. However, Contender bears what he calls ‘authority-based accountability’: we can hold Contender accountable in a robust sense because they furnished Goon with a bad purpose or goal. To evaluate Contender’s conduct, we

look at the actions of Goon (Bazargan-Forward, 2022, p. 36). In this way, we make sense of individual accountability in contexts of shared agency.

Bazargan-Forward considers his account less radical than Roth's and, as a result of the similarities between Roth's view and my account, less radical than reason transmission. The difference between Bazargan-Forward's and Roth's account is that the first claims that the reasons of another simply determine the purpose or end of my conduct in some cases. This is to say that my actions become a means to achieve your ends, which are determined by your reasons. On this account, it is purposes or goals that are transmitted from agent to agent via a division of agential labour. Conversely, on Roth's account, your reasons can be invoked to explain my conduct if we are working towards a common goal. On my account, when we act together and follow a shared plan, your reasons are transmitted to me via that plan. I argued that your reasons can motivate me to act when I follow a plan you created. While Roth and I state that the reasons of another can be my explanatory or motivating reasons, Bazargan-Forward is more cautious in claiming that it is simply your ends that I am acting on, not your reasons.

Bazargan-Forward criticises Roth's account of reasons and, by extension, my account, on the ground that when we have a reason to do something, that reason should serve as a premise in our practical reasoning (Schroeder, 2007; Way, 2015). On this classic view, we respond to reasons that we possess (i.e., have access to in our minds) and this reasoning allows us to rationalise our actions. The alleged problem that accounts such as Roth's and mine face is that when reasons are transmitted to us by someone else, we might not have access to them and the features that make those reasons favourable. The alleged risk is that, if we do not possess or know those reasons, they cannot serve as premises in our practical reasoning or

rationalise our actions. Since we cannot see what makes those reasons favourable to us, these transmitted reasons are not reasons at all, or so the objection goes.³⁶

This criticism is based on Davidson's influential account of acting for a reason. According to Davidson, a primary reason for an agent to do something consists of a pro-attitude of the agent towards actions with certain, desirable properties paired with a belief that an action has the properties in question (Davidson, 1963, p. 687). These primary reasons Davidson describes are explanatory reasons since they purport to explain why agents do what they do. This is also a causal view of reasons since they not only explain what agents do but also cause the agents to act in different ways (Davidson, 1963). In other words, there is a causal connection, for Davidson, between the agent desiring certain properties, believing that an action has those properties, and acting on the basis of those beliefs and desires.

On my reason transmission view and Roth's view of entitlement, this causal connection is seemingly broken because the agent reasons are transmitted to might not have the relevant pro-attitudes and desires toward the properties of the actions they are performing. In cases of acting together, there is a division of agential labour between two parties. One agent or group weighs reasons against one another, considers properties of different courses of action, believes that an action has those properties and then delegates the performance of that action to another agent. The latter allegedly does not have a pro-attitude towards actions with that property or a belief that what they are doing has that property, because the relevant deliberation was done by someone else. In this case, the objection goes, the reasons had by the deliberator do not explain what the executor is doing, or do not motivate the executor to act, because the executor does not see what makes those reasons desirable. The causal connection between the agent desiring

³⁶ This is the criticism Bazargan-Forward raises against Roth's account (Bazargan-Forward, 2022, pp. 54-57). Again, considering the similarities between my account and Roth's, I take that this criticism extends to my view as well.

certain properties, believing that an action has those properties, and acting is broken because the deliberating and the acting are done by two different people or groups.

The Davidsonian objection above would be valid if agents always needed to have reasons at the forefront of their minds for their actions to be based on or caused by those reasons. Even in simple cases of individual practical reasoning, agents do not always have reasons in their minds when they are acting. Yet, I argue, they act on those reasons. This is also valid in the interpersonal case, where multiple agents act together. I start by analysing diachronic intrapersonal action. I then show how the same principles apply to cases of diachronic interpersonal action. This demonstrates that we do not always represent reasons in our mind when we act on them, yet those reasons guide us and explain what we do.

Gemma wants to find a new job. She sees a job advert that ticks all her boxes – the job is remote, with a higher pay than her current role and great benefits - and wants to apply for it. At t_0 , Gemma engages in this deliberation and evaluates the pros and cons of staying in her current role and applying for the new job. She sees value in applying for the new job and devises a plan to do so. The plan entails writing up her job application the next weekend. She sets a phone reminder to do so on the relevant day. At t_1 , when the following weekend comes around and a reminder pops up on her phone screen, Gemma does not remember why she decided to apply for this job. She has forgotten the deliberation of the previous week and currently has no idea why this new role was desirable to her the week prior. Nevertheless, she acts on this plan made at t_0 . If she so planned, she thinks, she must have had a good reason to do so. She, then, writes and sends her job application. Does Gemma at t_1 act on the reasons that prompted her to plan to apply for the job at t_0 ?

According to Davidson, she does not. That is because at t_1 Gemma does not have the pro-attitude and desire associated with acting for a reason in her mind. At t_1 , she does not remember

what makes the action of applying for the new job desirable. In other words, she does not strictly possess the reason for which she is acting. So, the causal chain between her desires, beliefs and actions is seemingly broken.

However, Davidson's view that we need to have reasons at the forefront of our minds for them to explain our actions is too strong. We cannot expect agents to always recall the properties of actions they are acting on for them to explain their conduct or motivate them to act. It often happens that we make plans in the past that we execute in the future. When the moment comes for us to act, we might not remember the reasons why we planned to do something. Diachronic planning agency can be aided by forming plans, which allow us to organise our conduct to execute complex actions over time. If we have pro-attitudes and desires towards certain properties or actions, and we form a plan to act accordingly, that plan becomes an instrument for us to bring about the state of affairs that has those properties we value. We can deliberate and settle the matter on what to do and create a plan of how to bring about our preferred course of action at t_0 . When it is time for us to act on that plan at t_1 , we do not have to deliberate on what to do again: we can simply execute that plan to bring about our desired state of affairs. When we act on the plan at t_1 , we might not remember the reasons that prompted us to act. However, we do not have to consider them again or have them at the forefront of our minds for us to act on them. Since we planned with those reasons in mind, we can trust in our previous planning and, by following it, we can act on the reasons that prompted us to create that plan in the first place.

The Davidsonian causal connection between reasons and actions is preserved in cases of diachronic intrapersonal planning agency. When we have a reason to do something and form a plan to bring about the state of affairs that has those properties we value, a strong causal connection is formed between that reason, our plan and our future actions. The plan exists exclusively because we have a certain reason for action. If I had not been motivated by that

reason, I would not have formed the related plan. This plan exists because of that specific reason. There is then a causal relation between my reason and the correlated plan: that reason caused me to plan like that. Similarly, Gemma planned to apply for the new job at t_0 because that job has all the properties she values in a role. Since writing up job applications is a complex task that requires time to be executed appropriately, Gemma created a plan with relevant steps that would allow her to write up the job application. She designed this plan to bring about the state of affairs where she applies for the desired job. This plan of writing up her job application at t_1 exists exclusively because she has a reason to seek the new role. If she had not been motivated by that job being remote, offering a higher salary and so forth, she would not have formed the plan to apply for it. This plan to write up her job application at t_1 exists exclusively because she has reasons to go for that role. When Gemma follows her plan at t_1 , the reasons she had to apply for the new job explain what she is doing because that plan is an instrument to bring about the action that has the properties she desires. This is so even though she does not remember what those reasons were. Perhaps, Gemma has a bad memory. However, she is still acting for that reason when she follows the plan created to allow her to do so. The causal connection between reasons and actions is still present because the plan helps her to respond to her previous reason. Since plans allow us to coordinate our conduct to bring about complex states of affairs, they enable us to act on our reasons in the present or the future. They, then, ‘store’ the reasons they were created for, which we can act on in the future.

There are then cases when we do not represent in our minds the reasons we are acting on. However, those reasons can still rationalise our actions. Roth brings up a similar point when he discusses cases of acting on prior intentions. Sometimes, he argues, a consideration that functions as an explanatory reason for action is not actively represented in our mind at the time of action (Roth, 2027, p. 77). Once we decide to do something and plan accordingly, we do not always have to keep in mind our reasons to be motivated by them. Sometimes, we just do what

we planned without giving further consideration to the reasons why we planned. After all, as Roth puts it, “It is cognitively burdensome to have to keep track of all the reasons I may have for whatever it is that I am doing” (Roth, 2017, p. 78; Bratman 1987; Harman, 1986). The points above, then, demonstrate that reasons do not always need to be at the forefront of our minds for them to rationalise our conduct.

What about cases of diachronic interpersonal agency? Can the reasons of another explain and rationalise what we do if we do not possess them? Yes, they can. In cases of more agents acting together, the causal connection between desires, certain properties of things and actions stretches between different agents but is still maintained. To explain how this works, let us go back to Roth’s case of the two agents driving together to get downtown (Roth, 2017, p.82). In this scenario, both agents desire to go to the conference downtown. Agent 1, the navigator, believes that the action of turning left at the next exit has the property of getting them downtown faster. So, Agent 1 instructs Agent 2, the driver, to turn left. Agent 2, in this case, is acting on the desire they both have to go to the conference downtown, and Agent 1’s belief that turning left at the next exit will get them there faster. However, Agent 2 does not have direct access to that belief. Is Agent 2 acting on the reason Agent 1 has to turn left at the next exit? Davidson would say no since Agent 2 does not have the pro-attitude and desire associated with acting for a reason in her mind.

I explained, though, that we do not always need to have reasons in our minds for them to rationalise our actions. In cases such as the one Roth describes, shared plans allow agents to carry out complex actions with the help of a division of labour. If multiple agents are creating a shared plan of action towards a common goal, not all of them will have the properties of actions they are performing in their minds. In instances of shared agency, one agent might deliberate regarding the best course of action that serves a goal they share with others, where that deliberation contributes towards the creation of a shared plan. For every agent involved in

the shared activity, that plan becomes an instrument to bring about common goals and some properties the agents value. If Agent 1 deliberates on what they should all do and communicates the results of that deliberation in the form of a plan to Agent 2, Agent 2 does not have to deliberate again when they act on the plan. They can follow that plan to bring about the state of affairs they both desire. When Agent 2 acts on the shared plan, they might not know the exact reasons for Agent 1 to form that plan, however, they do not have to. Agent 2 can trust Agent 1's planning and, by following it, they can act on the reasons that prompted Agent 1 to create the share plan, which serves a goal they both share.

Similarly to diachronic intrapersonal agency, the shared plan above exists because the agents have reasons for actions and shared goals. When agents have a reason to do something together and form a related plan of action, a causal connection forms between those reasons, the plan and future actions. The plan only exists because the agents had reasons to do something together. If they were not motivated by those reasons, the plan would not have been crafted. The difference between intrapersonal and interpersonal agency is that the person who plans and the person who acts on the plan are not the same agent. However, since they are acting together towards a goal they share, those reasons for action can still motivate all agents since the plan exists to enable them to act on those reasons. The causal connection between the agent's desires, properties of actions and actions is preserved when different time-slices of the same agent plan at t_0 and execute the plan at t_1 . This allows agents to act on reasons they have now in the future and enables them to carry out complex actions over time. Similarly, some actions require the concerted effort of multiple agents to be executed. Shared plans, put in place to act on reasons agents have, enable agents to act together towards common goals and act on the reasons they had to create that plan, even though the properties or features of certain actions are not accessible to all agents involved in the shared activity.

So, it might be the case that, in acting together, some agents do not have direct access to the reasons on which they are acting. This happens when other agents oversee the planning of the shared activity, whilst they execute the plan. Do those reasons possessed by other agents serve as premises in the practical reasons of the ‘executors’ of the shared plan? Yes, if we are open to a more flexible approach regarding premises in practical reasoning. When I deliberate on what to do, there is a chain of premises that I follow in my practical reasoning. I might think that “I desire to go to the conference downtown”, that “driving will get me downtown faster” and that “turning left here allows me to avoid traffic”.³⁷ I can make a plan based on taking the car and following a set itinerary to get downtown faster. When the time comes, I can follow the instructions to turn left without necessarily thinking about why I must do so. The plan I made allows me to avoid deliberating on the matter again and simply follow the instructions I previously set for myself. “Turn left”, which is the content of my plan of action, allows me to act based on a previous premise I considered, which was “turn left here to avoid traffic”.

A similar chain of premises can be shared, where different sets of premises are distributed amongst different agents. The agents involved in the shared activity might all want to “go to the conference downtown together”, think that “carpooling is the best way to get there” and that “Agent 1 is the navigator” whilst “Agent 2 is the driver”. Agent 1, as navigator, plans the car journey and thinks that “turning left will allow them to avoid traffic”. They, then, instruct Agent 2, the driver, to turn left. The proposition Agent 2 considers in their mind is “Turn left”. By acting on that instruction, they can act based on a previous plan that, despite being crafted by someone else, serves a goal they have. This plan enables Agent 2 to serve a function in the shared activity, the function of driving them and their friend downtown, without having to reopen deliberation on what their partner already decided for them. “Turn left” becomes Agent 2’s premise in their practical reasoning and guides them on what to do next. Agent 1’s premise,

³⁷ All sentences in quotation marks, “...”, are premises in our mind.

“Turn left here to avoid traffic” is reflected in the instruction to turn left, since turning left is a way of bringing about the state of affairs to get downtown faster by avoiding traffic. Agent 2 can learn the features and properties of the instruction of turning left by asking Agent 1 to disclose those features. And this is how shared plans create a chain of premises in different agents’ practical reasoning.

I demonstrated that Davidson’s view is too strong when it comes to agents having to possess the reasons they act on. I defended transmission of reasons as an alternative approach. Before I conclude this section, I make another point to defend the Roth-style view of reasoning from criticism. Bazargan-Forward claims that his authority-based accountability approach differs from Roth’s view in that it does not claim anything as far-fetched as transmitting reasons to others (Bazargan-Forward, 2022, p. 55). However, his approach is highly compatible with Roth’s account and, in my opinion, can be considered an extension of Roth’s view.

Firstly, Roth states that I am partially entitled to your reasons in the case of driving to the conference downtown when your considerations serve an *end*, a *purpose*, that we share, which is to get downtown (Roth, 2017). When the agents acting together have a common goal, that they are trying to achieve through a division of labour where one acts and the other deliberates, the executor of the joint action is entitled to the reasons of the deliberator. Bazargan-Forward’s work also hinges on the concept of a common goal. In his view, goals are transferred from agent to agent. When agents agree to work together to achieve those goals and divide up the required labour amongst them, we can appeal to the deliberator’s reasons to make sense of the executor’s protected reasons. And we can appeal to the executor’s conduct to evaluate the deliberator’s accountability. In Bazargan-Forward’s view, then, a common goal and a division of labour are part of what we need to evaluate someone’s moral conduct based on the actions of the one who is executing their orders. If a common goal, a division of labour, shared agency and the execution of someone else’s plan are present, why does this not give rise to the kind of

entitlement to reasons Roth favours? If the criticism of the classic view of reasoning (Davidson's view) I presented above is convincing, then the cases Bazargan-Forward has in mind – with a common goal, a division of labour and shared agency – perfectly fit the bill of Roth's entitlement to your reasons account. If someone else's reasons might explain or motivate my actions, and I argued that it is, then it is not necessary to be conservative and claim that only purposes are transmitted from agent to agent. If reasons are robust elements that rationalise our conduct and guide us, and if it is possible to transmit them from agent to agent, I see no need to appeal to purposes in either attributing accountability or explaining actions in shared agency. I would rather appeal to someone else's reasons in rationalising my conduct than a purpose, as that reason would provide a complete account of my intentional actions, which includes the features someone else saw in that reason.

Secondly and lastly, Roth's and Bazargan-Forward's points are similar but with a different focus. Roth focuses on rationalising the explanation of joint actions. Bazargan-Forward is interested in how to make sense of an individual agent's accountability when what that agent does is performed with someone else. The focus of their work is different; however, the analysis is overall the same in that it purports to shed light on some difficult nuances of shared activity. It might well be that furnishing someone with an end is enough to make an agent accountable for something they are doing together. Perhaps, though, we need something more than just a purpose we are furnished with to explain our intentional actions. Perhaps we need to rationalise our actions in terms of someone else's reasons rather than someone else's ends. An account such as Roth's or mine is adequate to rationalise our actions in joint activities. If we, then, need to also establish agents' accountability in contexts of joint actions, it is possible to tack Bazargan-Forward's account onto something like Roth's account to complete this moral evaluation of conduct. That is why I see Bazargan-Forward's authority-based accountability as an extension of Roth's entitlement view.

5.1.2 Hierarchical Shared Agency and Group Hierarchy

How do agents act with others under the direction of authority? What is it to do something together in hierarchical settings? In this section, I analyse hierarchical shared agency, that is, shared agency characterised by authority relations amongst participants, which is pervasive in our sociality. Think of the former British Prime Minister, Boris Johnson, ordering the UK population to stay at home to reduce the spread of Covid-19. Or consider Greenpeace's board of directors hiring a spokesperson to defend the company's decision to build a new £ 45m ship. These are cases of hierarchical shared agency, namely, instances of agents doing something together as a result of the authority telling them to do so. Hierarchical shared agency can occur within the context of organised groups or institutions, but it does not have to. An important distinction needs to be made between hierarchical shared agency (agents acting together under the authority) and group hierarchy (the organisational structure of groups). Hierarchical shared agency and group hierarchy often go together, as members of hierarchical groups can act together under the authority's direction. They can, however, come apart, as coalitions of agents who do not belong to pre-existing, structured groups can also act together hierarchically (e.g., mobs following a leader). The mark of hierarchical shared activity (within groups or not) is a transmission of reasons from one or more agents to other agents. The mark of group hierarchy is agents occupying status roles with status functions in groups - where transmission of reasons occurs against this framework of roles and functions.

I start by defining and analysing hierarchical shared agency. I then move on to explain group hierarchy.

Hierarchical shared agency (HSA, hereafter) refers to agents acting together under the direction of an authority. A case of HSA would see participants following an authority that guides them and tells them what to do, and it is under its direction that they bring about common

goals. For a shared activity to be hierarchical, one or more of the agents involved must have power over the others and coordinate their actions. The agents in authority in cases of HSA are practical authorities³⁸ with the power to influence other people's behaviour. These authorities are put in charge of the organisational and decisional aspects of a shared activity. They coordinate the actions of other agents - which directly respond to them - and issue directives they must follow. The other agents involved in this shared activity will, then, be in a subordinate position to the authority that directs their actions. When they do as they are told, they participate in a shared activity with the authority issuing orders or directives.

The kind of practical authority I have in mind for cases of HSA is the authority to direct other agents' behaviour. I will agree that these authorities provide others with reasons to behave the way they are telling them to behave, and these reasons need to be taken as action-guiding by others (Ehrenberg, 2011). The authority's reasons then motivate our behaviour. The sufficient condition for acting hierarchically with others is that one or more agents succeed in making the others act for their reason. That is, the sufficient condition of HSA is that one or more agents successfully transmit their reasons to others. These reasons are transmitted in the ways analysed in the previous section (5.1.1).

Transmission of reasons from one agent to another does not occur automatically. Some conditions need to be present for one to successfully transmit one's reasons to another. In section 5.2, I present a six-step account that highlights sufficient conditions for an authority's reasons to be transmitted to its subjects. Here, I emphasise that there can be deviance cases where that transmission is broken. One way in which the transmission of reasons is broken is by reopening deliberation on a previously accepted plan of action. For instance, if I have a

³⁸ Practical authority is the kind of authority Joseph Raz (1979) is interested in. It mainly has a command-compliance structure, where someone in authority issues orders or directives and the subjects comply with them. This is in contrast with theoretical authorities, whose main role is to give people advice that they might or might not follow. For more on the distinction between theoretical and practical authorities, see Ehrenberg (2011); Hurd (1991); Green (2004) and Shapiro (2002).

reason to save money on food and plan to cook all my meals this week, that plan of action stores the reason it was created for. When I discharge that plan, I act on my reason to save money on food. However, after a few days, I forgot why I planned to cook all my meals this week. I reopen deliberation on the matter and decide to cook all my meals for the remainder of the week because I have fresh ingredients in the fridge that I need to use. In this case, after I deliberated again, I ended up with the same plan to cook all my meals at home. However, I now have this plan for a different reason. It is because I do not want my produce to go to waste that I now intend to cook my meals rather than to save money, which was my original reason. In this case, reasons are not transmitted from my original plan because I reopened deliberation and have a different reason to do the same thing. It might also happen that I forget why I made the plan to cook all my meals, reopen deliberation on it and end up intending to do so for the same reason as before (i.e., to save money). In this case, the transmission of reasons from my previous plan fails because I am only coincidentally acting on the same reason when I follow the plan I made at the beginning of the week. There will certainly be more cases where transmission of reasons fails, but I only have space to discuss one.

I now define group hierarchy. Group hierarchy pertains to the structure and organisation of groups, which can vary widely, ranging from simple two-agent collectives to mid-size companies, to large-scale corporations and nation-states. Despite the variation in size and levels of organisation, most groups have an (even minimal) internal structure that ensures their persistence over time. This structure is characterised by group members occupying different status roles with status functions attached to them.

Searle (1995) introduced the notion of status function to indicate the functions that objects, people or states of affairs have in social transactions partially due to those functions being collectively accepted by members of a specific society. Money, for instance, indicates a function – to serve as a form of payment and, generally, a medium of exchange – that some

objects such as pieces of (special) paper, metal and e-currency can perform. Various kinds of items have served as currency throughout the centuries (e.g., livestock, bones, seashells). What they all have in common is that, regardless of the material they are made of members of, agents regard them as the kind of objects that can be used as mediums of exchange. This function, attached to specific objects, is collectively accepted by members of societies and this collective acceptance is what enables objects to be regarded as money, and money to perform the functions it does.

When a status function is attributed to agents, it is referred to as a *status role*, which indicates the roles that agents can occupy with their related duties and responsibilities, such as being a teacher, a parent, or the president of the students' association (Searle, 1995; Ludwig, 2017). The role of 'teacher' indicates a function – passing knowledge onto pupils – that a person who meets specific criteria, such as having the relevant degree and training, can perform in an educational setting. Society collectively accepts that schools hire agents who meet those criteria to perform the function of transmitting knowledge. Someone who did not meet the relevant criteria would not have the status of teacher and would not be regarded as such by the community. Due to their conventional nature, status roles and functions need to be collectively recognised and accepted by group members to be operative. Once group members recognise and accept the status roles, agents occupying those roles acquire - *qua* role holders -specific duties and responsibilities. They are, then, expected to perform as their role requires and held to relevant standards of behaviour by other group members.

Group hierarchy is then characterised by a structure of status roles and functions attached to its members. We can say that agents are part of a hierarchical group when one person or a sub-group occupies a status role within the group that puts one in authority over others – call it authoritative status function - and this authoritative status function is acknowledged by other group members. The functions that agents or sub-groups in authority perform vary from role

to role. Agents with authoritative status functions, though, are usually responsible for directing and coordinating the conduct of others in some capacity. A headmaster has the status function of managing a school and its teachers; a CEO has the status function of managing a company and a university professor has the status function of teaching and supervising students. They all have ‘subordinates’ that rely on them to know what tasks to perform. The agents that have status roles are appointed to perform the functions that are attached to those roles, either formally or informally. For instance, one might be formally employed as head of marketing by a specific company as a result of an interview process that selects the best candidate for the job. Informally, one might find oneself in charge of a task after performing it for a prolonged time.

I am interested in the analysis of nation-states as large-scale, hierarchical groups of agents who act together towards common goals. In these groups, agents in different roles have access to different pieces of information. For this reason, asymmetry of information might occur and, when it does, it would affect the collective acceptance of status roles and functions in larger groups. I now explain what asymmetry of information is before going back to the discussion on hierarchical shared agency and group hierarchy.

Asymmetry of information occurs when some agents in the group have and hold access to information that is not possessed by or accessible to other group members. There are different pieces of information that some sub-groups of a hierarchical group might not have access to. One of them regards the creation or the specific content of shared plans.

Some agents or sub-groups might not know everything there is to know about shared plans. Asymmetry in information is often the result of the asymmetry of power in a group. Those who have the authoritative status roles of planning for their group might be the sole creators of the group’s shared plans. Other group members might not know how the plan was

created or do not have access to information that underlies the decisions made in selecting a particular plan. They might only be aware of the part they play in the plan, whilst being in the dark concerning other parts of the plan. In these cases, agents or sub-groups with authoritative status roles in a group have the function to dispense information to their subjects and, as a result, they have the power to choose what pieces of information to share with them and when. This kind of asymmetry of information can be negative, as not everyone will have knowledge of all the aspects of shared plans and, as a consequence, not everyone would be able to assess the adequacy or morality of the plan. However, it can also be positive because planning, especially the social planning that societies require, is time-consuming and requires significant cognitive effort. Not every citizen would be able to engage in social planning and having an authority that does that on their behalf can be advantageous.

Another piece of information that might not be accessible to some group members concerns the identity of other group members. This applies more specifically to large-scale groups, where not everyone knows all other group members or who the members with authoritative status roles are. Political societies are an excellent example of this kind of asymmetry in information. Citizens cannot know the identity or status role of all the other group members, yet they need to coordinate their actions with other people for social order to obtain. Political authorities provide plans that enable citizens to coordinate with one another without knowing the role of all the other group members. What citizens know in this case is that everyone else is supposed to follow the same set of laws and rules provided by the members with authoritative status roles. The coordination provided by those in authority gives them the ability to operate and act with others in societies without knowing all the other group members.

So, there can be several pieces of information to which members of large-scale groups might not have access. For this reason, asymmetry in information may be an obstacle to the collective acceptance of status roles and functions in large groups. According to Ludwig,

collective acceptance of status roles amounts to “enough members of the relevant community having appropriate generalised we-intentions with respect to action plans involving interacting with the agent with the role” (Ludwig, 2020, p. 186). I take this to mean that agents have roles in structured groups only if enough group members have intentions to engage with others in joint actions where certain agents play those roles. The intentions at play here are we-intentions on the part of the group members that, when directed at an agent, give rise to collective acceptance of that agent having a role. Ludwig’s account requires that enough members of the relevant community have these appropriate, generalised we-intentions towards agents in status roles for collective acceptance to work.

In large-scale groups, the kind of collective acceptance of status roles Ludwig describes seems difficult because of the asymmetry of information previously described. Some group members might not know who has certain status roles and what those roles entail. How can they accept something they do not know?

We should abandon the presupposition that agents can collectively accept status roles only if they know them. We should also abandon the assumption that we need enough group members to collectively accept status roles and functions. It is not plausible to have widespread collective intentions of joint acceptance in large-scale groups due to the high number of participants and the asymmetries of information.

A more fitting model of collective acceptance for large-scale groups would be a distributive model, according to which knowledge of the group’s roles and commitment is distributed among group members. This model involves the distribution of information that goes in different directions following the chain of hierarchy in the group. There can be top-down vertical distribution of information from those who are high in power, the operative members, to the other members or sub-groups who are lower down in the group ranks. This

vertical distribution can be bottom-up when people who are lower down in the group hierarchy inform the operative members of what goes on in the lower ranges of the group. Distribution of information might also be horizontal, as people who are in the same group rank can pass the knowledge onto one another. Group members are then disposed to accept status roles and functions, and they come to accept them upon acquiring pieces of information from other group members.

Furthermore, it is not necessary that enough group members collectively accept status roles. We only need relevant members to accept them. Not every status role would be pertinent to every group member. Do the lawyers in a company's legal department need to accept the status roles of the graphic designers in the company for the designers' status roles to exist and be effective? Perhaps only the graphic designers' line manager, the hiring committee and the board of executives need to acknowledge and accept the designers' status roles for them to be operative. The lawyers might accept the graphic designers' status roles if their own roles require them to be in contact with the designers at some point.

This distributive model of collective acceptance of status roles is not incompatible with Ludwig's account.³⁹ It specifies how information may be distributed in large-scale groups and how the knowledge gaps can be filled in for group members to collectively accept status roles. I now reprise my discussion on HSA and group hierarchy.

Stable groups maintain their structured network of status roles and functions whether they act hierarchically or not. Status role holders have functions to perform and are expected to do so by fellow group members. However, there can be moments in time in which the group is not engaged in any hierarchical shared activity without losing its structure. If the University Student Union's pizza joint closes for the summer term, all its activities are temporarily paused

³⁹ Ludwig, as discussed in conversation, agrees that agents being disposed to accept status roles is compatible with his account of collective acceptance of roles.

and, for the summer months, its employees and employers are not engaged in any hierarchical shared activity *qua* employees or employers. The pizza joint's hierarchical structure of status roles still exists – the line managers, the cooks, the servers, and the delivery drivers do not lose their jobs in the summer– but they do not engage in joint activities for that period.

Groups (and their status roles) can persist over time whether their members are always engaged in a (hierarchical) shared activity or not.⁴⁰ Similarly, a hierarchical shared agency might occur outside the context of structured groups. Agents can act hierarchically independently of their membership to any pre-existing group.

Consider this case.⁴¹

Blind Beach Rescue: A man is drowning in the sea. A beachgoer, call her S, has a reason to save the drowning man and knows that to do so a concerted effort from multiple agents is required. S quickly plans the man's rescue to respond to her reason to save him and starts shouting directives at three bystanders, B1, B2 and B3. These parties have never met before. They do what S tells them to do and act on her plan to save the drowning man. One thing to note is that B1, B2 and B3 all happen to be very short-sighted and, as a result, they are unable to properly see faces or figure out who is talking to them. Coincidentally, they all forget to pick up their specs from their bags before sprinting into action. However, S shouts very specific instructions and that, coupled with the blurred shapes they see, is sufficient for them to execute her plan and act on her reason. Prompted by S's instructions, B1 grabs a lifebuoy, B2 swims towards shore with the man, and B3 calls an

⁴⁰ This will not be the case for every group that pauses its activities, of course. If members of a book club decide to take a break for a month, and no one tries to contact the others to resume the group's meetings after the month has passed, they might – perhaps unintentionally – let the book club die slowly. In that case, the structure of the group would cease to exist after a while, presumably after the group members stop acknowledging and accepting the group's status roles.

⁴¹ Similar rescue cases have been widely employed in the literature, especially concerning the duty to rescue. See Douglas (2016), Ripstein (2000) and Smith (1990).

ambulance. Without seeing each other and being unaware of the other agents' actions, they manage to successfully rescue the man, thus discharging S's plan and acting on her reason.

Blind Beach Rescue is a shared activity. S and the three bystanders participate in a rescue effort together. They play their part in the joint activity of rescuing the drowning man by carrying out specific actions that, combined, bring about the goal of saving the man. This is also a case of *hierarchical* shared activity. B1, B2 and B2 – despite being unable to properly see – are responsive to S's instructions and acting as they do because S is telling them to do so. S is exercising power over them and coordinating their actions. The execution of her directives and her coordination are key to the success of their shared activity.

Furthermore, in *Blind Beach Rescue*, a transmission of reasons occurs from S to the bystanders. S has a reason to save the drowning man and realises, as we supposed, that such reason can only be fulfilled with a concerted effort from multiple agents. S, then, creates a rescue plan and starts shouting directives at the bystanders, thus telling them what to do. The bystanders act on S's plan, and they all manage to rescue the drowning man. S's plan stores the reasons S had to create it (i.e., a reason to save the drowning man, and reasons to direct people in ways that are apt to save him), and when the others act on them, they 'tap' into S's reasons for action. S took charge in that situation and the transmission of her reasons was successful when the bystanders started doing as she said. Note that the bystanders might have a (moral) reason to save the drowning man in this case. However, we can suppose that S also had other reasons, reasons to do certain things in a specific order to execute the rescue, such as a reason for B1 to keep the drowning man's head above water, a reason for B2 to call an ambulance and so forth. These reasons also get transmitted and make the hierarchical shared activity possible with S's coordination. The bystanders do not know exactly what S's reasons are. However, by

acting on her plans and directions, they indirectly act on her reasons and S's directions motivate them to act accordingly.

This HSA is not backed up by a pre-existing group. Before the rescue, the agents involved had not met before, and they would not see each other or act together again after this brief encounter. Agents can form pop-up groups that quickly come in and out of existence. However, for these temporary groups to arise, agents need to at least acknowledge each other's presence, be minimally responsive to each other's actions and avoid hindering the others for the success of the shared activity. In *Blind Beach Rescue*, each bystander is only responsive to S's directions. They are not wearing glasses, and this makes it so that they cannot make out the other agents they are acting with. They just happen to do as they are told without getting in each other's way. In short, the bystanders have no idea who they are acting with or if there is anyone besides S involved in this joint endeavour. If there cannot be a group – even a temporary one – without agents acknowledging and being minimally responsive to one another, then there is no group in *Blind Beach Rescue*.

Furthermore, it seems that there are no status roles and functions in this case. As we said above, the agents involved (apart from S) are not acknowledging and accepting each other as part of the same joint activity. The bystanders can only see blurred shapes and are unaware of who else is acting with them. They only respond to the instruction of someone they cannot properly see. They lack relevant information to be able to understand who else is involved in the rescue and what functions they are performing. Also, they are not mutually responsive to one another (only to S). Without collective acceptance on the part of the agents, status roles and functions cannot arise (Ludwig, 2020). For this reason, there are no status roles and functions in *Blind Beach Rescue*.

Blind Beach Rescue tells us that, despite the absence of pre-existing groups and status roles, hierarchical shared activity is still possible. So long as someone coordinates other people's actions and successfully makes others act on their reasons, HSA can occur. And we do have these elements in this case.

Hierarchical shared agency also occurs in the context of structured groups whose members occupy status roles, with status functions attached to them. In the context of groups, especially large groups, the transmission of reasons is a necessary but not sufficient condition for the group to act hierarchically. To carry out their activities and reach their goals, groups need a structure that guarantees their functioning over time. In one-off small cases of joint activity, the transmission of reasons is sufficient for HSA because the shared activity the agents are involved in is usually completed immediately or has a short lifespan. When a big group is involved – or when a shared activity is complex and requires agents acting together over time – the transmission of reasons on the part of the authority needs to be accompanied by something else. Each agent needs to occupy a role, with assigned duties and responsibilities, and the authority might also have to implement a system of threats or incentives for agents to carry out their portion of a shared plan. When groups are involved in hierarchical activity we need status roles and functions to define duties, responsibilities and expectations in the group and ensure the group and its activities can persist over time. Transmission of reasons from one or more agents to other agents occurs against this backdrop of roles and functions in groups. By transmitting reasons via intentions and plans, group members with authoritative status roles can direct and coordinate other people's behaviour towards common goals. This is how multiple agents manage to achieve complex goals and work together towards them in groups.

5.1.3 Nation-states as hierarchical groups

In this section, I analyse the hierarchical structure of nation-states as large-scale groups, with a specific focus on the status roles we can encounter in such groups. A comprehensive analysis of nation-states as group agents goes beyond the scope of this chapter and would be fit for another project. Here, largely following List and Pettit (2006), Ludwig (2017) and Shapiro (2011), I give a general overview of what the state's (and its government's) structure looks like if we frame it in terms of status roles and functions.

List and Pettit define groups as those “sets of individuals who are networked with each other in a way that matters to them or others and that affects their behaviour or that of others” (List and Pettit, 2006, p. 85). Members of a book club, for instance, are part of a group where agents are networked with other group members via their membership to the club and their interest in reading. Their behaviour is going to matter to and affect other group members (at least during the club meetings).

Similarly, nation-states are groups formed by many individuals who are connected due to them operating in the same legal jurisdiction or institutional organisation. Agents who live in the same state are networked with each other due to being residents or citizens of that state. Their thought and actions are also connected in relevant ways because they respond to the same state and abide by the same laws issued by the state. As large-scale hierarchical groups, states are highly organised to ensure the success of their plans and the smooth functioning of society as an institutional machine. We can identify three main sub-groups of the state as a large-scale hierarchical group, lawmakers, legal officials and citizens. These three sub-groups comprise specific status roles and functions assigned to group members. Let us analyse these sub-groups in turn.

The lawmakers, or legislators, are a sub-group of the state whose function is to create, update, and revise the laws, directives and rules adopted by the state. The lawmakers acquire their status roles and functions through a formal procedure where they are appointed to perform their roles. Lawmakers intend that their subjects (i.e., the citizens and residents of the state) act in specific ways, and they create plans for their subjects to follow, where these plans contain those intentions that the citizens act in certain ways (Shapiro, 2014, p. 266). According to Shapiro, an important part of the authority's role is to provide plans. More on this plan-providing function in section 5.2. Citizens, for the most part, do not create legal and political plans by themselves.⁴² These plans come from the authority. By creating and providing plans, the lawmakers perform the function of creating or modifying laws, policies, and directives for all the group members. Laws and policies are instructions that must be followed, while legal plans are a way of packaging those instructions for people to act on them.

Lawmakers do not exclusively devise plans for citizens or residents of the state to abide by. They also create plans that regulate powers and how they should be exercised. For instance, they can create plans that regulate their law-making functions by giving themselves rules that explain how to create laws and what falls in the realm of their powers. We can, then, distinguish between two main types of plans devised by lawmakers, let us call them general and procedural plans. General plans are those everyone in the group needs to comply with, as they coordinate and regulate group members' social conduct. Some general plans tell agents what to do, while others can also give rights and powers. Procedural plans are sub-group-related plans. They apply to certain sub-groups to regulate their functions, whether those are law-making functions or functions related to the enforcement of the law in the case of legal officials.

⁴² It is worth noting that sometimes citizens might rely on social norms to establish what to do when some laws are largely unenforced or unfollowed. Trespassing laws in some countries, for instance, might be unenforced and unfollowed for the most part. Citizens can appeal to social norms to decide whether it is appropriate to trespass in certain situations. Social norms might then be incorporated into plans of action in those cases where some laws are not taken as action-guiding.

The second main sub-group of states includes legal officials, namely, those agents whose primary task is to enforce the law. This category includes but is not limited to, police officers of all ranks, court officers (i.e., judges, magistrates, arbitrators, prosecutors, attorneys and so forth), and correctional institutions' personnel that deal with the punishment and treatment of criminals and offenders. These categories of officers have the status role of enforcing the law, whose functions include deterring, punishing, and rehabilitating those who break the law (Hess and Orthmann, 2008). Legal officials might also in some cases create laws and plans for others to follow. Judges are generally able to refine laws and can create new ones based on legal precedents by following decisions made in earlier cases by other judges. This is known as case law or judge-made law.

Legal officials need to follow procedural plans in carrying out their functions because those plans give them instructions on how to correctly enforce the law. Procedural plans created for legal officials also delimit the realm of their powers by instructing them on what they can and cannot do. A procedural plan for police officers, for instance, would tell them how to approach, address and treat offenders. That plan might specify that police officers can physically restrain someone (e.g., using handcuffs) during an arrest without causing any physical harm unless necessary, such as in cases of self-defence or to protect hostages. Legal officials, qua citizens or residents of the state, also need to abide by general plans, as any group member and citizen must do.

The last main sub-group we find in political societies includes citizens and, in general, residents of the state.⁴³ Every group member, including lawmakers and legal officials, belongs

⁴³ When I talk about the state's subjects, I refer to a broad category of agents that does not just include citizens. Not every member of this sub-group is going to have citizenship of the state. Some agents have the status of residents. Some other agents might only be in the state's territory for a short time, such as those who transit in the state's territory. So, when I refer to this sub-group, I refer to citizens, residents and those who transit in the state's territory, who all have to comply with the law.

to this sub-group. The citizens' main status role is to abide by authoritative plans.⁴⁴ Political authorities create and provide legal plans whose purpose is to coordinate the group members' actions, create social order and structure social life. Citizens rely on authorities to do the social planning for them because, as we will see in 5.2, it is more practical and efficient. In turn, they must follow those authoritative plans to achieve the coordination and order needed in political societies. The plans that citizens follow are general plans, such as 'Pay taxes by April every year', 'Renew your driver's license every 10 years', 'Only drive vehicles that comply with such environmental standards', and so forth. General plans constitute laws and directives authority intends their subjects to adopt to regulate their conduct. Some citizens will also belong to other sub-groups, for they can be officers of the law or lawmakers. In that case, they must comply with both general and sub-group-specific procedural plans.

Ludwig makes an important distinction between 'agent status roles' and 'subject status roles' (Ludwig, 2020, p. 187). Agent status roles are collectively accepted by their holders and enough members of the community. Subject status roles are assigned to agents regardless of whether they accept them. Agents with lawmakers or legal official status roles belong to the first category. They are appointed to their roles via a procedure during which they accept to take on those roles. Conversely, prisoners of war have a subject status role because they do not accept their role, which is assigned to them by others. Citizens, for Ludwig, have a hybrid status role because citizenship is assigned to them at birth regardless of whether they accept it, making that of citizen a subject status role. However, citizens are presumed to have collectively accepted their roles when they reach the age of majority, after which their status roles become agent status roles. In our classification of the three sub-groups, the citizen sub-group is an

⁴⁴ There are some cases in which citizens participate in the creation of the law. Some US states have citizens-initiated *referenda*, which allow them to place new legislation on a ballot for popular vote. Citizens can then introduce a law or constitutional amendment through a petition in this case. For a reflection on citizens-initiated *referenda* in Australia see Williams and Chin (2000). For an analysis of the future of local government in Europe see Denters and Klimovsky (2017) and for an empirical assessment of this type of *referenda* see Qvortrup (2013).

example of a sub-group with hybrid status roles, since citizens get assigned their status role at birth and residents participate in the collective acceptance of their roles.

One last remark on the citizen status role before I conclude this section. The main function attached to the citizen status role is to execute plans provided by the authority. Executing plans is an important function. Plans are crucial to execute complex projects in the present and the future. Creating plans, which is something the lawmakers are tasked to do, is the first step towards fulfilling common goals in political societies. However, the creation of plans alone is not enough for the state's members to reach their goals. It is essential that they follow them and put the intentions behind plans into action.

Authoritative plans are formed by intentions that track the goals authorities want members of the state to achieve. Authorities are, then, the creators of plans that are shared with other group members. Since creating plans alone is not enough to fulfil them, group members must implement authoritative plans. This is where the citizens' sub-group comes into play. Without the citizens putting authoritative plans into action, those plans could not be fulfilled. The citizens' contribution, in the form of executing plans, is essential to the very functioning of the state. Without them fulfilling authoritative plans, namely, without them abiding by the law, there would be no social coordination and social life would not run smoothly. If nobody abided by authoritative plans, those plans would serve no purpose and the functions of political authorities themselves – to organise and coordinate social life – could not be achieved. In this sense, authorities depend on the citizens' sub-group for their plans to be successful.

Ludwig recognises the role and prominence of the subjects of legal systems. The subjects execute the authority's plans in legal systems and have rights and duties that stem from their role. In Ludwig's words "the law has internal authority over its target subjects, those for whom it is designed" (2020, p. 200). This does justice to the fact that legal systems do not merely

comprise legal officials. They also incorporate the citizenry that, with legal officials and lawmakers, constitutes the polis. Lawmakers and legal officials have important status roles, but they perform their roles with the citizens in mind: they represent the citizens by acting as proxy agents for them.⁴⁵

In the next section, I present an account of how political authorities plan for their subjects and delve deeper into the lawmakers' status function of creating plans for others to follow.

5.2 An account of hierarchical shared agency for nation-states

In this section, I analyse how political authorities plan for their subjects. I present a six-step account of hierarchical shared agency. It explains how political authorities create plans and how their subjects incorporate them into their practical reasoning. This account of how authorities plan is also an account of sufficient conditions for an authority's reasons to be transmitted to its subjects. Transmission of reasons does not occur automatically, and reasons will not be transmitted if the conditions I highlight later in this section are not present.

In section 5.1, I defined hierarchical shared agency as an instance of agents acting together under the direction of an authority that transmits its reasons to them. By transmitting its reasons to other agents, the authority motivates them to act towards common goals and influences their behaviour. I also said that an authority transmits its reasons to others via intentions and plans. I think of plans and intentions as metaphorical repositories or vessels that store the reasons they were created to respond to. Since reasons are transmitted from agent to agent via intentions and plans, we need an account of how authorities plan for their subjects.

⁴⁵ Cf. Shapiro to whom legal activity is a case of shared agency that sees legal officials creating and applying the law together. Shapiro's category of legal officials corresponds to the lawmakers' and the legal officials' subgroups in my account. In Shapiro's account, legal officials are full-fledged participants of the legal shared activity, whereas citizens are not because they are not involved in the creation and application of the law (Shapiro, 2014, p. 288).

In what follows, I sketch an account of authoritative planning that is modelled on political authority. This account accommodates what Shapiro (2014) calls ‘massively shared agency’, namely, large-scale shared activities, but might also fit cases of acting together in smaller hierarchical groups. This account explains how authorities create plans to coordinate other agents’ behaviour and how agents incorporate authoritative plans into their practical reasoning.

In Chapter 4, I presented a six-step account of shared practical reasoning for egalitarian shared agency. That account explained how agents act together as equals, without authority relations among them. Can this account of shared practical reasoning generalise to incorporate cases of hierarchical shared agency? The account I presented in Chapter 4 cannot generalise to explain cases where agents act under the authority’s direction. Some of the intentions involved in egalitarian shared practical reasoning cannot be possessed by many in large-scale groups.

As a reminder, the six-step account of egalitarian shared practical reasoning unfolds as follows:

- 1) The group has shared goals to achieve.
- 2) Group members form end-related intentions toward the shared goal (e.g., I intend to do X with Y).
- 3) Group members form means-related intentions connected to their end-related intentions (I intend to do my part of X).
- 4) A shared plan is formed for the group to act on.
- 5) By default, group members must treat the shared plan (and its related reasons) as exclusionary.
- 6) 1-5 is common knowledge between participants

According to this account, individuals acting together have two main mental states that contribute to shared activities: end-related and means-related intentions. Agents acting together

have common goals to achieve and can direct their actions toward bringing about those goals. Participants in the shared activity have intentions that are directed toward an outcome they want to bring about with others. I argued that when acting together individuals intend to do something with others, that is, they intend to do X with Y. These are end-related intentions each participant in a shared activity has. The content of these intentions is the entire action or state of affairs agents are bringing about with others. In small-scale settings, individuals can come together and devise plans of action. They will then have a general sense of what they want to achieve with others and how to go about achieving it. They can also form means-related intentions to carry out their portion of the shared activity (i.e., intentions to do one's part of X, where X is the entire shared activity).

In large-scale groups, agents involved in shared activities can direct their actions toward bringing about common goals and form appropriate means-related intentions. However, not many agents in large groups would be able to form end-related intentions to do something with other agents. Asymmetries of power and information that are typical of large-scale groups prevent most group members from forming intentions that have the entire shared activity as their object. Agents with authoritative status functions who exercise power over others in big groups often create plans for their subjects to follow. Authorities intend that others act in certain ways and can provide them with plans of action to fulfil those intentions. In nation-states, lawmakers create plans in the form of laws that the citizens have to follow. Since most ordinary citizens do not participate in the creation of the law, they do not have access to all the details of authoritative plans.

Power imbalances between group members in nation-states cause asymmetries of information in the group. Those who did not create laws or plans, might not have enough details about the content of shared plans created by authorities. Citizens, for instance, might not know how some plans were created and what purpose they serve. They might only be aware of some

parts of bigger plans that have many ramifications in political societies. For instance, citizens might know that there is a plan in place for them to wear face masks inside public buildings during a global pandemic, but they might not be aware of the bigger plan and strategy their government is putting in place to combat the pandemic. They can form intentions to wear face masks as a contribution towards the goal of fighting the pandemic, but it is unlikely that they can form intentions that are directed at the totality of the government's strategy to combat the pandemic.

So, we need an account of hierarchical shared agency for large-scale groups that can explain how members of these groups act together without many of them having end-related intentions with the totality of the group's shared activities as their content. The account of hierarchical shared agency for big groups I present below is modelled on Shapiro's account of what he calls "shared intentional activity with authority" (Shapiro, 2014, p. 266), which is the equivalent of what I call 'hierarchical shared agency' in large-scale groups.

Shapiro (2014) expanded on Bratman's (1999a) account of egalitarian shared agency to include cases of agents acting with authorities (Shapiro, 2014). According to Bratman's classic account of egalitarian shared agency, agents participate in a shared activity, J, if and only if (a) they both intend that they J; (b) they both intend that they J because the other intends that they J; (c) they both intend that they J in accordance with sub-plans that mesh; and (d) all of this is common knowledge between them (Bratman, 1999a, pp. 100-101). I expand on this account below.

The paradigmatic cases analysed by this model are those in which two people are engaged in simple activities (e.g., going for a walk together, or painting a house together). The participants are equals and the joint activities they carry out do not involve hierarchy. (b) and (c) are two important constraints to shared agency that matter in our analysis. Firstly, regarding

(b), it is not enough for shared agency that individuals both intend that they J (i.e., that both intend that they do something together). For the shared activity to proceed smoothly, Bratman thinks that participants need to intend that they do something in virtue of the other participant's intention that they do something. This is the interlocking intentions constraint. For instance, I need to intend that we paint the house in virtue of your intention that we paint the house (and vice versa). If you do not intend that we paint the house, then my intention that we paint the house will either be abandoned when I realise you do not intend that as well, or my intention will not be part of a shared activity. Secondly and lastly, interlocking intentions are still not enough for successful shared activities. The participants also need to have sub-plans that mesh (c), where this entails following a shared plan coherently, compromising in the face of disagreement and not hindering other agents' contributions to the joint activity. For Bratman, the meshing sub-plans constraint also ensures that the participants are mutually responsive to each other's actions and intentions.

I largely followed Bratman in sketching my account of egalitarian, shared agency in Chapter 4. There, I did not include an interlocking intentions constraint, namely, I did not argue that agents need to intend that they do something in virtue of the other agents' intentions that they do something. However, means-related intentions to do one's part of X are generally accompanied by a belief that a certain task is my part to play in X, where X is a shared plan. This intention and belief might also be accompanied by an expectation that other agents in the group are going to play their part in the joint enterprise.

When you decide to do something together with others, you and your group members might also have mutual expectations⁴⁶ that other group members are going to play their part.

⁴⁶ By mutual expectations here I mean that group members in some cases expect other group members to do their bit. Another reading of mutual expectations sees them as common beliefs with the higher order beliefs that each believes that all believe that all others do their bit. I do not espouse this reading here.

These mutual expectations are thought to anchor the joint activity and get it off the ground in some cases.⁴⁷ When individuals expect their group members to collaborate, they are (ideally) less prone to free-riding and more prone to intending to play their part in the shared activity knowing that they are not the only ones working toward the common goal. The expectation that others are going to play their part also plays a role in the coordination of thought and action in the group. If you expect your group members to perform as decided, you know what they are going to do and when, and that is going to guide your action when you do your part in the joint activity. The expectation that others are going to do their bit is not a necessary component of a shared activity. You can do your part despite lacking the expectation that others will play their part. While this may be unwise, some might say irrational, we can imagine situations in which one group member is very committed to the group's cause and will do their bit whilst being unsure that other group members are going to play their part. Expectations that other group members are going to play their part, then, might or might not arise in contexts of joint actions but their absence does not negatively affect how everyone's thought and action is directed. Mutual expectations that everyone is going to contribute to the common project are essential for shared activity in Bratman's eyes, given that each intends that they (the group) J (does something together) by way of others intending the same thing. This interlocking system of individual intentions is not necessary in my account. Expectations about other people's intentions and actions might arise out of the belief that we're playing a part in a shared activity and our intention to do so, but they are not essential for individuals to intend their bit of the common project and for them to act on that intention.

However, my account of egalitarian shared agency includes a version of Bratman's meshing sub-plans constraint. Each individual's means-related intention to play one's part in a shared activity X relates to other agents' intentions to play their part in X. Different parts of the

⁴⁷ Such as in Gilbert's notion of joint commitment (1999).

same shared activity carried out by different agents need to mesh in a way that allows them to complete the shared activity together. Participants in a small, egalitarian shared activity also tend to be mutually responsive to each other's actions and intentions in completing their shared project.

Like my account of egalitarian, shared agency, Bratman's classic account of shared agency struggles to accommodate cases of hierarchical shared activity in big groups. Shapiro argues that the main obstacles to adapting Bratman's account of shared agency to cases of hierarchical shared activity are the interlocking of intentions constraint and the meshing sub-plans constraint. Regarding interlocking of intentions, Shapiro notes that when agents act together by way of following an authority that tells them what to do, they might not always intend to do something by virtue of the other participants' intentions to do the same (Shapiro 2014, pp. 270-271). It is more likely that subordinates only intend to do what the authority says without tracking what other agents intend. This is an example of vertical interlocking of intentions when subordinates interlock their intentions with the authority's intentions and not the other participants in a shared activity (Shapiro 2014, pp. 267-268). In cases of hierarchical shared activity, agents intend to do what the authority intends that they do, but they might not intend to do that because other agents that are not the authority also intend to do the same.

Regarding the meshing sub-plans constraint, Shapiro argues for an indirect mesh of the participants' intentions in hierarchical shared activity. In large-scale contexts, the authority designs a plan of action for other agents to follow and ensures that everyone's part of the shared plan is coherent and consistent with other sub-plans. The authority is responsible for other agents' meeting the meshing sub-plans constraint. Having meshing sub-plans means that agents involved in a shared activity compromise in the face of disagreement, do not hinder other agents' contributions to the shared plan and, ultimately, are mutually responsive to each other's actions and intentions. Agents' ability to compromise when needed and to not get in each

other's way is mediated by the authority in cases of hierarchical shared agency. The authority creates a plan of action and assigns sub-plans to other agents. The authority then oversees the execution of the hierarchical shared activity and the solution to disagreements, should they arise, whilst ensuring the coordination of everyone's actions. Other agents do not have to worry about coordination and solving disagreements themselves. Furthermore, they do not even need to be responsive to each other's actions and intentions. Given that each of them responds to the authority's intentions that they do something, their concern is to comply with those orders and, so long as they do that, they are carrying out their sub-plan of the main plan. Each agent can discharge their part of the plan despite being unaware of what the others' sub-plans entailed. They might not know what the authority is asking the others to do, but that is not necessary for them to complete their task. They only need to know their part of the plan because the authority's organisation and overseeing of the plan is enough to meet the coordination needs that meshing sub-plans require. It is the authority's activity that performs the function of the meshing sub-plans (Shapiro, 2014). Bratman's meshing sub-plans constraint does not apply to cases of hierarchical shared activity in large groups without a modification of what the constraint itself entails (i.e., from direct to indirect mesh).

For these reasons, Shapiro argues that a shared intentional activity has authority relations when there is an authority that 1) intends that the subjects take what she says as their sub-plans and ensures that they mesh. Moreover, 2) the subjects intend to take the authority's orders as their sub-plans and to be responsive to changes in those orders. (1) and (2) are also common knowledge between all the parties involved (Shapiro 2014, pp. 266-267). So, to accommodate for hierarchy in Bratman's account, Shapiro modifies the interlocking of intentions and the meshing sub-plans constraints Bratman adopted to explain egalitarian shared agency.

I adapt and modify Shapiro's theory to sketch an account of hierarchical shared agency in nation-states that incorporates the elements I discussed so far. According to this model,

members of large-scale groups, including the citizens and residents of nation-states, do not always share an intention to act together with other group members. They might intend to do something with others if they have enough information about the plan they are following, but this is not always the case in large-scale, hierarchical contexts. Participants in hierarchical, large-scale activities, however, share plans of action, which are created by the authority for them to follow. The authority is responsible for creating plans for its subjects and coordinating their actions.

The six-step account of hierarchical practical reasoning I propose unfolds as follows:

- 1* A nation-state has collective goals to achieve
- 2* The relevant political authority intends that its subjects do X, where X is a shared activity in the pursuit of a common goal
- 3* The authority creates plans (laws) for the subjects to follow (to complete X)
- 4* The subjects form means-related intentions to play their part in X because of the authority's intentions in 2
- 5* By default, the subjects must treat the authority's plans in 3 as exclusionary
- 6* 1*-5* is common knowledge between the parties involved

1* deals with the fact that groups, including nation-states, can have collective goals to achieve. These goals are collective in the sense that they pertain to the group, and their achievement is sought by the group via its members acting together. Among the collective goals political societies have, we can find creating and maintaining social order, solving civil or penal disputes among parties, coordinating group members, solving coordination problems, defending the state's territories and so forth.

In the case of large-scale groups, asymmetries of information make it the case that not every group member knows about the group's goals. Lawmakers, and perhaps legal officials, might have a complete picture of the collective goals to achieve and know the goals' contents. A shared goal in this case is the outcome a collective activity is directed to. Given that authorities organise their subjects' conduct via plans, authorities know the outcome the group's activities are directed to. Other group members, such as the authority's subjects, however, might be unaware of the group's specific goals. In large-scale groups, the distributive informational model we employed for collective acceptance of status roles and functions (section 5.1) works for acquiring knowledge of collective goals as well. Upon becoming members of the group, agents will be disposed to accept the group's goals. This is because membership in a group often entails working together with other group members towards common projects. So, with membership comes the disposition to collaborate with others, accept status roles and functions and act together to bring about common goals. Individuals will know about common goals upon acquiring information from their superiors in hierarchical groups. In nation-states, lawmakers might communicate collective goals to their subjects. Also, individuals can have a sense of what society's goals are (in the case of states) by looking into what laws lawmakers are passing. Laws aim to bring about collective goals by instructing people on what to do. A law that forbids driving gas vehicles in city centres on the first Sunday of every month to reduce carbon emissions, for instance, informs the subjects that environmental protection is something lawmakers want the group to engage in. The horizontal and vertical transmission of information in large-scale groups is, then, an important tool to spread knowledge of the group's goals. The subjects do not need to have all the details of the outcome the group wants to bring about. They need to have a general sense that what they are contributing to with their actions is directed at an outcome they are bringing about with others in their nation-state.

Political authorities (e.g., governments in nation-states) can intend that their subjects do something (2*). More precisely, authorities can intend that the group they exercise power over, authorities included, engage in a shared activity. These are end-related intentions that take the form of ‘I intend that they do X’ or ‘We intend that they do X’. These end-related intentions have the totality of the shared activity the group needs to carry out as their content. In smaller groups, authorities might intend in the present that their subjects do something and create a plan of action for them to act accordingly, either in the present or the future. In nation-states, some laws were promulgated in the past and are still active in the present. The government officials in charge might not actively intend that their subjects act in accordance with laws they did not promulgate, that is, they might not have the correspondent mental state in their mind. And authorities will not have corresponding intentions that their subjects do something for every law that has validity in their territory. Lawmakers will not know every law or their content. Lawmakers, however, directly intend that their subjects do something when they pass new laws and can intend that the subjects comply with the laws that apply in their territory, even ones that were promulgated before their mandate.

3* is about authorities’ ability to create plans that their subjects must or should follow. As we said, authorities can intend that their subjects do something. When they so intend, they can create plans for others to follow based on those intentions. As mentioned in section 5.1.2, authorities might create procedural and general plans for their groups. Procedural plans are specific to each sub-group of the main group. They regulate the functioning of sub-groups by establishing rules sub-groups members need to follow. Furthermore, procedural plans delineate the boundaries of someone’s power as they tell people what they can and cannot do. General plans are to be followed by everyone in the group, as they coordinate and regulate group members’ conduct. In nation-states, general plans are laws that tell citizens what to do or assign rights and duties to them. These laws also apply to legal officials and lawmakers, as they too

are citizens of the state. Lawmakers' planning activity is devised to coordinate and organise citizens' conduct to solve coordination problems and prisoner's dilemmas, among other things. States have the resources, infrastructure, power, and time to engage in this social planning that individuals do not have. The state's planning activity facilitates acting together in political societies. Authoritative plans, influence and direct the citizens' conduct in specific ways and, as I explain shortly, eliminate some options from individual practical reasoning.

As a result of the authority intending that they do X, the subjects can form means-related intentions to play their part in X (4*). These intentions are connected to the authority's intentions that they do X and are individual intentions in the group members' minds to do something as a result of the authority telling them what to do. The subjects in nation-states would most likely not be able to intend to do X with others, where X is the totality of the shared activity they are engaged in. They can, however, intend to play their part in X when they come into contact with authoritative plans, either after the authority communicates that the plans are in place or after observing other agents' behaviour.

Agents can form means-related intentions to play their part in the authority's plans after they accept them. Acceptance of the authority's plans is an individual mental state that encompasses the recognition on the part of an agent that the authority created plans (and sub-plans) that involve them.⁴⁸ Those plans were designed with the idea that group members have a part to play and that their contribution is required to discharge plans and fulfil the group goals. When agents see that certain authoritative plans involve them, they can incorporate those plans

⁴⁸ Acceptance of authoritative plans will not be universal in nation-states. We only need a sufficiently high number of members to accept authoritative shared plans, and I leave this vague intentionally. When a good number of group members accept authoritative plans, individual actions are well-coordinated, and things run somewhat smoothly. Large-scale groups can afford some group members free riding and disregarding laws. When the number of people who accept and follow authoritative plans is higher than the number of free riders, coordination and social order are guaranteed in the group. Problems arise when the number of free riders is so significant that it either trumps the number of law-abiding citizens or it is so high that it severely disrupts societal activities. Lawmakers need to keep the balance right between law-abiding citizens and free riders, and they have different means of doing that, such as backing up laws with sanctions and adding incentives to encourage people to respect the rules.

into their practical reasoning. Incorporating the authority's plans into one's practical reasoning means that one will act on the authority's plans and take them into account in designing their own plans. Group members do not need to know all the details about the authority's plans to be able to act on them and discharge the shared plans with other group members. They only need to know their portion of the shared plan. Authorities oversee the design of plans and the coordination of people's thoughts and actions so that individuals can focus on their portion of the plan. In the case of citizens, the only thing they need to know is that must follow the law, which specifies their portion of the shared plan. Citizens and members of large-scale groups in general, do not need to have a complete picture of the group to act together with others and on the authority's intentions. It is sufficient that they recognise that, in abiding by the law, they are part of an institution where other people do the same thing to achieve common goals, such as solving coordination problems. Abiding by the law is intending to play one's part in society's shared plans.

Before moving on to the next step, I elaborate on how agents incorporate the authority's plans into their practical reasoning. This process involves an active decision-making component. When one plans to take the bus to work, for instance, one makes decisions and sub-plans as to which bus would get to the office faster, what time one should take it to get to work on time and so forth. If there is an authoritative plan that requires bus passengers to purchase a ticket for their rides, and one believes there is such a plan, that authoritative plan is considered by the agent because her decisions are affected by it. The agent needs to factor in that she might be denied access to the bus without a ticket, which might result in her being late for work if another mode of transportation is not found on time. She might also consider trying her luck, hopping on without a ticket in the hope of not getting caught. Or she might decide to purchase the ticket before her ride as per the authority's plan. In deciding to purchase a ticket, the agent is actively incorporating the authority's plan into her reasoning. She needs to take the

bus to go to work and there are laws surrounding the use of buses which are going to affect her plans. In purchasing the ticket, the agent is plugging this authoritative plan into her own reasoning. Purchasing the ticket allows her to take the bus to go to work, which is what she planned to do. In purchasing the ticket, the agent is also accepting the authority's plan. Acceptance of authoritative plans means recognising that a plan involves you (or that it was created for you in this case) and deciding to adopt it. It is worth noting that acceptance does not have to always involve an elaborate decision-making process. Agents can accept authoritative plans quickly and automatically, as those plans can become part of their routine. Our commuter, for instance, might be used to buying bus tickets every time she goes to the office, and the process of accepting and incorporating authoritative plans into her reasoning becomes automatic for her and she might not need to revise her plan every time she takes the bus.

As previously stated, authorities can plan for their subjects and the subjects can intend to play their part in the authority's plans. By default, the subjects must treat authoritative plans as exclusionary (5*). I argued in previous chapters that plans in general, and their related intentions must be treated as exclusionary for them to successfully guide agents' conduct. Individuals can reopen deliberation if necessary. However, after they settle the matter of what to do and create a plan, the plan must be treated as exclusionary by default until more stringent reasons to do otherwise arise. Authoritative plans function the same way. In hierarchical groups, someone in authority is vested with the status role of creating plans for others. After the authority deliberates and settles the matter of what the group and its subjects should do, it creates plans that the subjects must treat as exclusionary. The difference between the authority case and egalitarian shared agency is that, in large-scale hierarchical groups, the majority of group members have little to no say in the creation of the group's shared plans, which are imposed on them from above. Practically, when individuals adopt authoritative plans and treat

them as exclusionary, those plans block their deliberation and prevent them from choosing otherwise in most cases. I will explore whether this impacts individual autonomy in section 5.3.

Moreover, all of this (1* to 5*) is common knowledge among the agents involved in hierarchical shared activities. By this, I mean that group members know that, when they abide by the law, they are playing their part in a plan put together by the authority for them to coordinate their actions with others. They do not need to know who the other group members are or all the details of the plan and the goals they are working towards to form means-related intentions to play their part in the authoritative plans. This common knowledge constraint only requires agents to be responsive to the authority and its plans.

Finally, this six-step account explains how authorities plan for their subjects, but it also contains the conditions needed for reasons to be transmitted from the authority to the subjects. Reasons are transmitted when agents act together towards common goals and when there is a division of labour in the group whereby one or more agents make a plan for the others to follow. In this case, when the authority creates plans to coordinate the subjects' conduct, the subjects, who share common goals with the authority of their nation-state, can act on the authority's reasons when they follow authoritative plans. I reprise this discussion in Chapter, 6, Section 6.1, when I compare my account of transmission of reasons with Roth's account of entitlement to someone else's reasons.

5.3 Individual Autonomy and Authoritative Plans

In this section, I analyse how the relationship between authorities and their subjects affects personal autonomy. I investigate whether abiding by the authority's laws and plans threatens the subjects' autonomy. I focus on treating the authority's plans as exclusionary. Does the fact

that individuals treat those plans as exclusionary in hierarchical settings render them less autonomous? It does not. The argument I present in this section builds on the one I proposed in Chapter 4 to show that shared plans in egalitarian contexts do not hinder personal autonomy. That argument answered a worry about shared plans and autonomy in egalitarian cases. The argument I make here raises and answers yet another worry that stems from the relationship between authorities and individuals.

The argument I presented in Chapter 4 addressed concerns related to acting autonomously in contexts of egalitarian shared agency, where no one involved in the shared action is in a position of authority over the others. Does the presence of authority in a group or a shared activity change the equation? Is personal autonomy affected when agents do as the authority says?

Prima facie, there is a worry for autonomy in the hierarchical case. Most members of large, hierarchical groups, as we noted in section 5.2, do not participate in the decision-making process required to create shared plans. These plans, made by those in power in the group, are imposed on other group members, who must follow them. In nation-states, lawmakers have the power to create plans in the form of laws that they and the other sub-groups must follow. Citizens are not usually involved in designing these plans and, as we saw in discussing asymmetries in information, they might not know all the details of those plans. Does this make them less autonomous when they abide by the authority's plans? It does not. Even in hierarchical contexts plans devised by an authority are useful instruments for agents to respond to reasons and they also allow people to make decisions despite the constraints imposed by the authority.

What I say below follows the arguments for the compatibility of individual or egalitarian, shared plans and personal autonomy. Plans, regardless of who created them, have a similar

structure and follow the same norms of practical rationality. Acting alone might seem an inherently different activity than acting on the authority's plans but, at the core, they are similar. They are both about following a plan and letting it guide and inform your actions. The difference lies in who created the plan. Sometimes we act on and follow plans we did not create and, while this is a useful time-saver tool, it is also necessary in some cases. Our view of autonomy accommodates the possibility of acting on the authority's plan and, in general, relying on others.

The argument for the compatibility of autonomy and authority, or the compatibility of authoritative plans and individual practical reasoning, rests on two points. Firstly, authoritative plans can help agents to respond to reasons (a Razian perspective), and secondly, agents can still decide for themselves even though they are following plans created by the authority. I will discuss them in turn.

In understanding how authorities, with their plans, can help their subjects respond to reasons, we need to consider the main difference between egalitarian plans and authoritative plans. In discussing egalitarian shared plans in Chapter 4, I noted that individuals often join groups whose reasons align with theirs and that this alignment of reasons partially explains how they retain their autonomy in acting with others. This is a rarer phenomenon in hierarchical groups. Some hierarchical groups, such as workplaces, certain clubs or associations can be voluntarily joined by their members. One can, for instance, voluntarily join a company one admires as an employee after a successful job interview. In such cases, hierarchical groups are voluntarily joined because individuals have goals in common with those groups (e.g., advancing the renewable energy field by working for a renewable energy company). When individuals already have goals that align with some of the group's goals, and those goals are better achieved by acting with others, individuals have reasons to join some groups. When agents act according to shared plans that group members in authority impose on them, the

reasons that stem from those plans are already their reasons. Those reasons are part of a hierarchical shared activity motivated by goals that group members already have in common. In these cases, the reasons of the group align with the reasons of the individual members who, as we supposed, voluntarily joined the group.

Other cases of hierarchical groups, though, are different. Nation-states are largely not the kind of hierarchical groups that can be easily joined at will. Most people are born in a state without the possibility of choosing otherwise. It can also be difficult, although possible, to leave one state to move to another. When individuals cannot choose their state, they find themselves to be members of groups whose goals or reasons differ from theirs. They might not even know those goals or reasons or care about them.

When agents are born into hierarchical groups they did not choose, the authority in the group and hierarchical plans they formulate can nonetheless help agents to respond to their reasons. This perspective comes from Raz (1979). As members of political societies, individuals have a reason to coordinate their actions with the actions of other group members and adjust their conduct accordingly. It would be difficult to pursue personal plans without this kind of coordination. They would hinder each other's plans to fulfil their own, with the consequence that pursuing individual goals would be difficult and costly. If one wants to purchase a house in a sought-after area, for instance, one needs to coordinate one's actions with other people's actions to succeed in that complex endeavour. And if there are no rules that regulate the house-selling market, rogue individuals might make it difficult for others to purchase properties. Large-scale interpersonal coordination, which is the hallmark of political societies and most large-scale hierarchical groups, is a tool to facilitate the fulfilment of individual goals and plans. Members of political societies have a reason to coordinate their conduct with that of other citizens. This coordination often enables them to respond to their other reasons, as it simplifies how they can act with rules and procedures. This is where

political authorities come into play. They take care of the organisational aspects of social life and provide their subjects with rules and regulations, in the form of 'ready-made' plans, to incorporate into their practical reasoning. In this sense, authority's plans are designed to facilitate individual practical reasoning, by eliminating the need for costly, large-scale coordination on the part of individuals at every turn.

This is an instrumental and indirect way in which authorities can help their subjects respond to their reasons. Individual reasons might not align with the authority's reasons and goals. Agents might also be indifferent to the authority's reasons. Nonetheless, they are assisted in pursuing their reasons and goals through the authority's plans and reasons. By complying with the law in nation-states, they can pursue personal projects without (ideally) being hindered by other agents. For instance, if one needs to go from A to B, and there are laws in place on how to behave behind the wheel, one can get to B by following those laws, thus reaching one's goals and responding to one's reasons. Authoritative plans provide a framework of acceptable behaviour in large-scale groups and purport to coordinate the group members' conduct. Even if agents are not interested in the group's goals or have different reasons than the group's reasons, authoritative plans allow them to respond to their reasons by providing rules of conduct to the group members. Note that this is compatible with the authority modelling the group's plans on its member's reasons, which is what Raz thinks legitimises authority (Raz, 1986). However, my account is also compatible with the authority creating plans that are based on reasons group members do not share. If those plans are adequate, and I present some adequacy conditions of plans in Chapter 6, they would still allow agents to respond to their reasons.

In nation-states and large-scale groups in general, individuals might come to identify with their group's reasons over time. This is problematic in some cases and there can be times when the group's reasons and individual reasons are not aligned. States may make promises

and create plans that they subsequently retract. Individuals who share the authority's goals and plans at t , might not share them at t_1 if there is a drastic change in those goals and plans. A state, for instance, might set the goal to reduce carbon emissions to comply with recent environmental regulations and its subjects, who are environmentally conscious, praise it for that. However, unbeknownst to them, the state then signs a lucrative deal with an oil company and fails to uphold the new environmental regulations. The citizens who previously praised the state for its environmentally conscious goals and plans are now not aligned with the state's goals and plans. They previously shared the state's goals and reasons, and the state's plans were a way for them to respond to their reasons, which aligned with the state's reasons. When things change and there is a misalignment between the citizens' reasons and the authority's reasons, we cannot say that the authority directly helps people to respond to reasons. If it does, it does it indirectly and instrumentally.

The second point of the argument for the compatibility of autonomy and authority is that agents can still decide for themselves when they follow authoritative plans. How? When agents are governed by someone else, the processes of individual or shared practical reasoning presented in Chapters 3 and 4 do not disappear. In hierarchical contexts, individuals still have personal or shared goals and can create plans to achieve them. However, agents need to incorporate the authority's practical reasoning into their own. The six-step account of hierarchical practical reasoning presented earlier explains the mechanism behind how we incorporate authority's plans into our plans. An example might help us understand those mechanisms further. I intend to buy a car and I proceed to form a plan and relative sub-plans that will enable me to fulfil that intention. The successful execution of this plan is threatened by the fact that I do not have enough money to buy the car. I consider two possibilities to overcome this obstacle. To acquire the money I need, I could either rob a bank or ask for a loan. Upon reflection, I realise that robbing a bank is not an option I should consider because it is

against the law of the state I'm subjected to. Therefore, I decide to ask for a loan instead. I then proceed with my plan and buy the car.

This simplified case explains how authority enters individual practical reasoning. Members of nation-states have a duty to abide by the law. They ought to consider the legal background of laws in their reasoning. This background of laws and authoritative plans filters out options that are, at least in theory, not admissible to them (because they are against the law) and it constrains and directs their intentions and choices.

Wolff (1970) would argue that these legal constraints, and the fact that authorities with their plans filter out certain options for us, still hinder personal autonomy. He might think that constraints that come from authoritative plans a) make it so that not every option is available to us to choose from when we make decisions, and that b) following a plan entirely made by someone else, does not allow us to exercise our capacity to decide for ourselves.

Claim (a) can be responded to by clarifying what we mean by deciding for ourselves. Deciding for ourselves does not entail that we must continuously deliberate and make decisions at every turn, or that we need to be the ones deciding every option that goes into our plans. We make and follow our life plans,⁴⁹ act based on our goals and follow our desires and passions, but this does not exclude, and cannot possibly exclude, relying on others. Concerning acting with others in egalitarian contexts, I said that we should abandon a conception of autonomy that excludes reliance on other agents (Chapter 4). Such a theory of autonomy confuses self-determination – autonomy as making your own decisions and responding to reasons – with complete self-reliance – autonomy as relying exclusively on yourself for every action or decision. This is also true, and even more relevant for agents as members of political societies. There are complexities associated with being members of large-scale groups. To achieve social

⁴⁹ Susan Dodds and Karen Jones (1989) call the capacity to envision and create a life plan for oneself 'dispositional autonomy'.

order, ensure the population's safety and promote justice among other things, individuals must coordinate their actions with that of other group members. This large-scale co-ordinational activity is bound to be costly and time-consuming. Had they to tend to this massive activity of coordination themselves, individuals would not be able to successfully fulfil their plans and respond to their reasons for action. Moreover, the limited cognitive abilities of rational agents would not support that level of organisation.

In such cases, individuals can defer complex choices and organisational issues to an institutional body that has the resources to manage large-scale coordination problems. Given that they cannot take care of social planning, agents can delegate part of their planning powers to the authority. That authority would create a set of laws whose function is to facilitate interpersonal coordination among many people, solve coordination problems and settle disputes that may arise between the subjects (Shapiro, 2014). They might not be the only institutional body capable of such complex large-scale coordination. A deeper geopolitical analysis might reveal that states and governments are not needed for that purpose.⁵⁰ However, the fact that most nations rely on states and governmental institutions for that kind of coordination and the fact that coordination is often mostly achieved by those states and governments, warrants an analysis of states as large-scale groups capable of coordinating people's actions.

In response to b), the worry that we do not exercise our capacity to decide for ourselves when we follow plans created by others, we frequently allow for the possibility of following someone else's plans. This happens when agents hire others for certain services, or when they are part of various hierarchical groups. One might, for instance, seek the services of a

⁵⁰ The case of Belgium is rather interesting, as the state has survived long periods without an official, elected government. In August 2020, Belgium broke its standing record for the longest period without an elected government (592 days as of August 2020) [<https://www.brusselstimes.com/124777/belgium-breaks-own-record-for-longest-period-without-government>].

nutritionist to learn how to eat more healthily. The nutritionist will then create a meal plan that one should follow to improve one's diet. In following the nutritionist's instructions, one is following a plan made by someone else. That plan is based on a goal that one already had, namely, eating more healthily, but the plan was entirely designed by someone other than the agent who is now executing it. This dynamic is at play in many instances where agents hire other people to do things for them or on their behalf. Another typical case of acting on other people's plans is the employment case, where employees act on plans devised by their employers. After signing their contracts, employees are part of a hierarchical group where they routinely play their part in plans they did not create.

When we seek other people's services or rely on someone else to plan certain aspects of our lives, we delegate tasks to other people. Delegation, the act of deciding which projects we can complete or tackle by ourselves and which projects we should give to someone else to complete on our behalf, is embedded in our societies. Delegation is a tool with which we can maximise our time and efforts to reach our goals effectively.

If we had to decide everything by ourselves, without relying on other agents to plan for us or do things on our behalf, we would not be able to complete all our projects and reach our goals. Delegation of our planning powers is a big part of our lives as group members, and a conception of autonomy that does not allow us to delegate our planning would not fit with how rational agents behave. If being autonomous does not allow acting on other people's plans, it would follow that we are, in fact, rarely autonomous, considering how heavily we rely on others to plan or act on our behalf. We have, then, two choices: we can either accept that there are many instances in which we lose our autonomy, in acting with others or in following someone else's plans, or we can redefine our view of autonomy to accommodate the complexity of our lives as social beings.

5.4 Summary

I analysed hierarchical shared agency, the phenomenon of agents acting together under the authority's direction. In section 5.1.1, I presented my account of reason transmission, compared it to similar approaches to reasons in the literature and defended it against criticism. In section 5.1.2, I distinguished between hierarchical shared agency and group hierarchy. Agents often act together under the authority's direction in the context of organised groups or institutions. However, it can be the case that coalitions of agents who do not belong to pre-existing, structured groups act together hierarchically (e.g., mobs following a leader). I argued that the mark of hierarchical shared activity (within groups or not) is a transmission of reasons from one or more agents to other agents. The mark of group hierarchy is agents occupying status roles with status functions in groups, where transmission of reasons occurs against this framework of roles and functions.

In section 5.1.3, I analysed the hierarchical structure of nation-states as large-scale groups. Following List and Pettit (2006), Ludwig (2017) and Shapiro (2011), I provided a general overview of what the state's structure looks like in terms of status roles and functions. I distinguished between three main sub-groups of nation-states, lawmakers, legal officials and citizens or subjects. This paved the way for the account of how political authorities plan for their subjects in the following section.

In section 5.2, I investigated how political authorities in nation-states plan for their subjects. I presented a six-step account of hierarchical shared agency that explained how political authorities create plans and how their subjects incorporate those plans into their practical reasoning. This analysis was crucial because I said before that authorities transmit their reasons to other agents via intentions and plans, which are metaphorical repositories of reasons. Since reasons are transmitted from agent to agent via intentions and plans, we needed

an account of how authorities plan for their subjects. Authorities create plans to coordinate other agents' behaviour. In turn, individuals can incorporate authoritative plans into their practical reasoning.

Finally, in section 5.3, I argued that autonomy and authority can be compatible. The argument for the compatibility of authoritative plans and individual practical reasoning rested on two points. Firstly, authoritative plans can help agents respond to reasons, and secondly, agents can still decide for themselves even though they are following plans created by the authority.

In Chapter 6, I turn to the legitimacy question and examine the possibility of legitimate political authorities. I put forward the delegation theory of the legitimacy of political authority, according to which political legitimacy is the product of large-scale, hierarchical agency.

Chapter 6 - The Delegation Theory (Part 1)

6.0 Introduction

In previous chapters, I resolved the *prima facie* tension between autonomy and shared agency. Resolving that tension was important to establish the possibility of legitimate authority.

I can now turn to the legitimacy question in this chapter, which focuses on the delegation theory of the legitimacy of political authority. This delegation-based account understands political legitimacy as a product of large-scale, hierarchical agency.

According to the delegation theory, political authorities are legitimate when they meet two conditions:

- a) When they provide their subjects with adequate plans that help them respond to reasons;
- b) When they give a basic public justification for their plans.

I argue that, when conditions (a) and (b) are met, the subjects ought to delegate their social planning powers to the authority.

The chapter is divided into three further sections. Section 6.1 deals with the first condition of legitimacy. It analyses the desideratum we want to find in authoritative plans (what makes them ‘adequate’). Section 6.2 explains why authorities need to provide a public justification for their plans and how they do so; it also argues that the duty to delegate on the part of the subjects is unlocked when the authority provides adequate plans that are publicly justified (Chapter 7 is dedicated to the problem of delegation). Section 6.3 offers concluding remarks.

6.1 Authoritative plans

The first condition that political authorities need to meet to be legitimate is to provide their subjects with adequate plans that help them respond to reasons. This first condition immediately presents us with an issue. What does it mean for authoritative plans to be adequate? This section delves deeper into this issue that arises out of the first condition of legitimacy. I now analyse the desideratum for adequate plans.

There is one, albeit multifaceted, desideratum that authoritative plans need to meet to pass the adequacy test. To be adequate, authoritative plans need to be sensitive to solving societal problems and, in doing so, they need to account for and respect all the relevant moral considerations. Let us tackle these elements in turn.

Plans created by political authorities need to be sensitive to societal problems, where these might be coordination problems, problems of creating and maintaining social justice, prisoners' dilemmas, and compensating people in need in cases where social justice has failed to accommodate them. These problems inevitably arise out of living together in large-scale groups such as political societies. We can expect local variations, but the types of problems are likely going to be similar from society to society. Why do we need the state to solve these types of problems? We partially addressed this question in previous chapters. Coordinating your conduct with that of others is an essential aspect of living and acting together (Locke, 1689). By its very nature, societal coordination involves many people, thus being more complex than any single individual can handle. It requires concerted efforts to be executed. Single individuals do not likely have the time, knowledge and resources to execute this social coordination. Institutional bodies have the time, knowledge and resources to do so. Given the importance of social coordination, individuals should rely on others to achieve it when they cannot do so themselves. So, individuals should rely on institutional bodies to achieve social coordination.

In current political societies, the state – embodied by a group of lawmakers - seems to be the default institutional body that takes care of the overarching organisational aspects of social life. This is not to say that social coordination cannot be achieved by other groups other than lawmakers or that it will not be achieved by other groups in the future (AI networks?). Given that the state is the default institutional body capable of such large-scale coordination, I will assume that it is the institutional body agents should defer to if they cannot achieve social coordination efficiently themselves considering current the state of the world.

So, if the need for solving societal problems causes individuals to rely on the state's resources and expertise to do so, and if the state operates – as I argued throughout the thesis – by creating plans for its subjects to follow, it is a necessary condition of legitimacy that the state's plans are well-suited to address those societal problems. Plans that do not adequately address societal problems are not going to coordinate and direct people's conduct as an attempt to solve those problems, which is the main (if not only) reason why we need the intervention of the state in this case.

However, it is not enough that authoritative plans are sensitive to societal problems. The authority might implement plans that track societal problems very well but are so far-fetched (i.e., unjust or immoral) that no rational individual would be willing to act on them. For instance, a plan to boost food production in a country that involved forcing minorities (of the population) to work long hours in the fields to meet high demands for food cannot be considered adequate. It might be that this plan tracks a problem that needs to be solved (i.e., low food production in the country), but it would be wrong to go through with it since it ignores important moral considerations (i.e., the welfare and rights of that population's minorities). Whatever problems a society is facing, its solutions need to incorporate and meet important moral constraints. The authority cannot impose plans that require its citizens to do immoral things, and authoritative plans should not violate the subjects' most fundamental human rights,

intend to cause emotional or physical harm to individuals if that is not done to prevent them from harming others and so forth.⁵¹

The desideratum that authoritative plans be adequate is, then, multifaceted because it includes considerations of efficiency and morality. We assumed that efficiency and morality are not separated. They can perhaps be disentangled as *critical considerations*, but they have to be co-realised in any acceptable solution. Plans are not to be considered adequate in our theory if they would be efficient at solving societal problems when the solutions they put forward involve the states or their subjects doing something immoral. Embracing a version of adequacy that ignores some important moral considerations would be misguided in our theory particularly because we have argued throughout that authorities, with their plans, can help the subjects respond to reasons. Since the subjects' pool of reasons includes moral reasons, authoritative plans should be sensitive to those to be able to help the subjects respond to reasons.

A question that arises from the discussion above is how demanding this account of authority's legitimacy is going to be. I base this account on a *scalar model* of the legitimacy of political authority, and this makes it less demanding than an account of political legitimacy could otherwise be. According to the scalar model, political legitimacy follows a gradient whereby an authority can be more or less legitimate depending on the plans it puts forward. Moreover, the subjects' reasons to respect the state's authority are graded in line with the degree of adequacy the authority's plans reach. Individual reasons to follow the state or respect its plans can be weaker when the authority routinely fails to produce adequate plans in some areas,

⁵¹ We can imagine that striking a balance between good – meaning well-functioning – plans and moral plans is not an easy task for states and that there are going to be many cases where the well-functioning feature and the moral feature of plans will come apart. Moreover, morality might look different in different states. One might also ask whether it even makes sense to talk about the morality of a state (Sorley, 1901) and its plans. I do not have the space to fully resolve this issue here. But I believe it is not unreasonable to ask states to at least refrain from violating the most fundamental human rights (Shue, 1996), and for their plans to reflect this effort.

or stronger when the authority has a history of putting forward consistently solid plans in most areas. The state would have different areas of operation, such as healthcare, education, welfare, external affairs and so forth for which it puts forward plans. The state might fail to produce adequate plans in the sphere of healthcare, for instance, but its plans might nonetheless be more successful in other areas. Suppose the state has to decide whether to fund certain drugs. It ends up making the wrong decision and since it failed to provide the right drug, many of its subjects die. If we endorsed the scalar view of the state's legitimacy, we would not say that the state ceases to be legitimate after this one failed plan. Such a state would still be legitimate in all areas, including healthcare, provided that its other plans are adequate. However, if the state consistently makes mistakes in the healthcare sphere, its subjects might have weaker reasons to abide by its laws on healthcare-related matters. They might still trust that the state is going to deliver adequate plans in other areas of legitimacy and therefore abide by all but healthcare-related laws. Over time, the state might become completely illegitimate on matters of healthcare, whilst still maintaining legitimacy in other spheres where its plans are adequate. If the state consistently fails to put forward adequate plans in all areas, it will become completely illegitimate over time.

Authorities might get things right on certain matters and make mistakes in others. If we espouse the scalar model of legitimacy, we will think that they lack legitimacy in some ways albeit not completely. We need to establish, though, how well-directed authoritative plans need to be at solving the societal problems we identified. Being clear on how successful or efficacious authoritative plans need to be is important to determine when the authority's operations fall too low on the scale of legitimacy.

We should reject a model of perfect efficacy whereby authoritative plans are adequate when they *are* solutions to societal problems. In this view, only those plans that succeed at solving problems are adequate and, as a result, only the authorities that come up with the best,

bullet-proof solutions to their subjects' issues are legitimate. We should reject this view on grounds of demandingness, as very few (if any) current states would be legitimate on it.⁵² We cannot expect states to get everything right. Things will not always go according to plan and lawmakers also incur human errors. We cannot demand of anyone that they get everything right, so we cannot demand the same of the state. This account of legitimacy is too implausible to have any successful practical application and so it needs to be rejected.

Another option would be to say that plans are adequate when the authority devises them to solve societal problems. We can interpret the term '*devise*' to mean that authoritative plans are not actually solutions to social problems, or that they are the best solutions. Through the creation of plans, the authority aims to solve societal problems, and this is separate from whether its plans are going to succeed at doing so.⁵³ All that is required from the state, in this case, is that it puts forward plans that track the issues that need to be resolved at that moment in that specific society and try to solve them. The authority is not going to be blamed or suddenly become illegitimate if its plans fail to deliver the expected results. This account considers the possibility of human errors and includes in the sphere of legitimate authorities those states that make an effort to create adequate plans, even though those plans might not succeed at reaching the common goals the society set.

On the face of it, the account of legitimacy above seems to make space for a variety of different possible legitimate political orders. The sphere of legitimacy here includes political authorities whose plans have different degrees of success: some fail to solve societal problems, others only solve certain aspects of those problems, while others address issues the right way and fulfil the goals the plans were created to achieve. But they all put forward plans that aim

⁵² Of course, it might well be argued that this area of political morality simply *is* very demanding and there is not much that a theory can do to avoid placing those demanding constraints on real states. I believe, though, that we can find a balance between how states realistically operate and how demanding a theory of legitimacy should be.

⁵³ I will insist on some degree of actual efficacy.

to be good solutions to certain problems. And it is compatible with this approach that some states are going to do this with a better success rate than others.⁵⁴

This is a complex view that combines the scalar model of legitimacy with the claim that good intentions, rather than results, are what matters. I accept the scalar model of legitimacy, but I think we need to reject the claim that we should only count good intentions for an authority to be legitimate. This claim lowers the threshold of legitimacy too much. All that this view requires for the authority to be legitimate is that the authority tries to create plans to solve societal problems. It does not matter in this view whether it succeeds or not. However, an effort to solve societal problems on the part of the authority regardless of its outcome is not enough. We need to consider some degree of expected efficacy in planning. Authoritative plans need to track the societal problems they were created to solve and put forward solutions that are expected to have a good degree of success. This requires surveying different options before landing on the one that, on a cost-benefit analysis, seems to have a greater chance of success. How much success or efficacy authoritative plans are expected to have will vary in practice, depending on the circumstance, the problems they are expected to solve and the urgency of those problems. In theory, we do not need to pinpoint exactly how efficacious or successful plans need to be. What matters is that the authority is expected to not simply try to put forward solutions to problems, but to choose solutions that it deems efficacious and to deliver some results (of course things might not always go according to plan).

To sum up, the account of political legitimacy I have in mind is one where authorities produce plans to solve societal problems and have a history of achieving good results. Their plans are adequate when they are chosen as solutions to societal problems for their expected

⁵⁴ This approach differs, for instance, from an early Rawlsian (1957; 1958) approach to political legitimacy according to which only those states that embody the ideals and principles of justice as fairness are legitimate political states.

chance of success. Adequate plans, then, need to have an expected degree of efficacy. Exactly how efficacious and successful they need to be will vary in practice. On this account, political legitimacy follows a scalar model whereby there are different areas of legitimacy and authorities can be more or less legitimate in those areas and as a whole depending on the plans they produce.

The last point I need to clarify in connection with the demandingness of political legitimacy deals with purity of intent in planning. Does the authority need to be motivated to devise plans that are tailored to that goal? Is it still legitimate if it has ulterior motives? I said earlier that the authority is legitimate when it devises plans to solve societal problems, and we can interpret this as the claim that authority aims to solve these problems. According to this interpretation, what the authority intends (the authority's motivation for doing something) matters, and it needs to be motivated to devise plans that are tailored to that goal. However, it might not be necessary that the authority is motivated to solve societal problems.

Suppose that a state is not concerned with its subjects' problems at all. The lawmakers' primary goal is to stay in power. They are aware that the only way for them to stay in power is to keep winning elections, and that to keep winning elections they need to help their subjects solve societal problems. So, they devise plans that will try to reach that goal. Here, the authority's motivation is not to help people solve problems. The authority does not plan with that in mind. It plans in such a way because it is motivated to remain in power. Does the authority need to be motivated to implement plans that successfully track societal problems to be legitimate? If we answer yes to this question, we risk excluding from the sphere of legitimate authorities those who implement the right kinds of plans for different reasons than helping subjects solve their problems. If the authority in the above example puts forward plans that track societal problems in the right way, does it matter that it did so to stay in power? I am

inclined to think that it does not matter what the authority's motivation is so long as its plans target societal problems with a certain degree of efficacy.⁵⁵

In the next section, I explore public reasoning on the part of the authority as the second condition of legitimacy on the delegation theory. Public reasoning can facilitate the transmission of reasons from the authority to the subject, as it would be a way for the subjects to know what the authority's decision-making process entails.

6.2 Public Justification on The Part of The Authority

In the last section, I analysed the first condition of the legitimacy of political authority. I argued that authorities are (partially) legitimate when the plans they impose on the subjects are adequate and help them respond to reasons. I clarified what I mean by adequate plans and how the subjects' intentional action is explained by their reasons and the authority's reasons. It is now time to turn to the second condition of legitimacy. Producing adequate plans that help the subjects respond to reasons is not enough. To be legitimate, authorities also need to provide a basic public justification for their plans. In this section, I explain what that public justification amounts to and how it works. I also address some concerns that arise regarding the compatibility of public justification with the concept of delegation and the notion of authority I endorse.

By saying that the authority needs to provide a basic public justification for its plans, I mean that the authority needs to explain what its plans are for and say something about their plausibility as solutions to societal problems. This is to reassure its subjects that the plans track

⁵⁵ It is worth noting that reliability of effort is also important. If the only way the authority can retain power is by solving societal problems, then it may matter less that it is motivated by power. The desire to preserve power and to solve societal problems are marching in step in this case. If, by contrast, there were other ways of retaining power that didn't involve solving societal problems, then this would be less satisfactory, because the authority might choose to reach the goal to stay in power in ways that do not include solving societal problems.

common goals the society has and that they are means to those ends (represented by the common goals). This, then, makes this second condition of legitimacy an epistemic condition that the authority justifies to its subjects the plans it is offering, which translates into epistemic access on the part of the subjects to the authority's reasoning.

What does this public justification look like in practice? Take the Covid-19 example. We assumed that the authority had a reason, R , to reduce the spread of the Covid-19 virus with a national lockdown. There is a fact that p that the spread of the Covid-19 virus can be decreased by 50% with a national lockdown and the authority has that reason as a result of knowing this fact. So, the authority creates a plan that instructs its citizens to stay at home. For this plan to be publicly justified, the authority needs to communicate what the plan is for and roughly explain how it is a plausible solution for a problem the society has at that moment. In this case, to publicly justify the national lockdown plan, the authority would have to say that the plan serves to contain the spread of a virus and that the national lockdown is a plausible solution to the spreading of the virus as it has a certain percentage of efficacy at reducing its spread. When the authority provides this kind of explanation, its national lockdown plan is publicly justified. When the authority passes a plan without that explanation, the plan is not publicly justified.

Why do we need this public justification of the authority's plans for the authority to be legitimate? There are two main things to be said about adding this epistemic requirement to our conditions of legitimacy for political authorities. Firstly, it is reasonable that individuals know why they are asked to do certain things and what certain plans are for. This is so in the case of demanding plans such as the national lockdown plan. If agents are required to stay at home for a long time – thereby sacrificing a considerable portion of their freedom of movement and social life – it is not unreasonable that they would want an explanation as to why that kind of plan is imposed on them. This applies to less restrictive plans as well. If there is a plan in place that instructs agents to drive below a certain speed, they might expect some sort of

explanation as to why that is the case. After all, why should they not be allowed to drive faster, thus reaching their destination sooner? Agents might want to know why it is the case that they cannot drive faster in certain areas upon being presented with a plan about speed limits. Secondly, knowing the explanation behind authoritative plans might be beneficial for compliance. Some agents might be more willing to follow authoritative plans if they can see the rationale behind them and know why they are asked to behave in that way.

So, knowing the rationale behind authoritative plans and having an idea – albeit vague – of what they are for can help the subjects understand whether the first condition of legitimacy has been met. One could object that it is beneficial for the authority to justify its plans for matters of compliance and clarity as to why it is asking its subjects to do something. However, this does not mean that public justification of the authority's operations is necessary for the authority's legitimacy. I believe that in a context where the authority expects the subjects to treat its laws and plans as exclusionary, the subjects must have a way of checking that the authority is doing what it is supposed to do. Exclusionary plans alter individual deliberation by making it so that agents take those plans at face value without personally assessing what to do. This is unproblematic in those cases where the authority's plans are adequate (i.e., when they meet our first condition of legitimacy). But when the authority's plans are inadequate, second-order reasons to reconsider those plans arise for the subjects. If the subjects do not have a way to check on the authority's directives when needed, how can they know whether they should reconsider the directives or obey them? The subjects are unable to assess the authority's operations if the authority's laws and directives are to be treated as exclusionary. A way for them to know that authoritative plans are adequate is for the authority to publicly justify what is doing. When the authority communicates the reasons behind its decisions and plans, the subjects can verify that those plans are adequate in a straightforward way to do so that does not require too much work or involvement on their part. This public justification condition can be

more or less strong. In some cases, the authority might provide a case-by-case justification of all of its plans, whilst in other cases, it might just be prepared to justify its actions should that be required. I am more inclined towards an account of public justification where the authority makes provision to justify its operations and makes relevant information about this justification available to the subjects who can access it when needed.

Public justification on the part of the authority matters because what the authority does can have a big impact on our autonomy. Having a mechanism in place to check that the authority is operating appropriately is a way to honour Wolff's anarchist worry that both voluntarist and non-voluntarist obligations to obey authority can endanger our capacities to make decisions and respond to reasons. With public reasoning, we have a way of knowing which authoritative plans hinder our autonomy and which do not. Furthermore, public reasoning gives our account of political legitimacy an advantage over those accounts that lack an 'escape clause', a way for the subjects to express their dissent or do otherwise. Raz's account, as we saw in Chapter 2, does not allow for the subjects to re-open deliberation freely if needed when they take authoritative reasons as exclusionary. It only allows them to discard authoritative directives in those cases decided by the authority in advance. The advantage of public reasoning is that it puts matters into the subjects' hands: it is up to them to decide whether authoritative directives need to be reassessed and discarded. If the authority is putting forward plans that do not allow the subject to respond to reasons, second-order reasons to do otherwise would arise for the subjects. These second-order reasons would clash with the authority's second-order reasons which require the subjects to ignore their first-order reasons that go against the authoritative directive. When the subjects perceive this clash of second-order reasons, they can check on the authority's operations via public reasoning. If they realise the authority is putting forward adequate plans, they can do otherwise and discard the directives.

Another objection to public reasoning that might arise is that we do not need the authority to publicly justify its plans because we could achieve the same result when the subjects have sufficient reason to trust the authority and the plans it puts forward. To this, I respond that trust needs to be built and earned and that regular public justification of plans might be required to initially build and maintain trust in the authority over time. In general, the preparedness of the state to justify itself to subordinates is a condition for them to reasonably trust the state's plans. Moreover, there will always be new entrants to the state as individuals reach maturity. These individuals will not have been around long enough for authority to have earned their trust. As a result, there has to be, in effect, a continuous flow of public justifications in a state.

We said that the subjects have a duty to delegate to the authority which entails following the authority's plans when they are presented to them. If we think that there is a need for the epistemic requirement that the authority publicly justifies its plans – and we said above that there is - this means that the delegation condition is not going to be immediately satisfied when the first condition of legitimacy is satisfied. To put it differently, the duty to delegate to the authority will not be unlocked until the subjects know the justification behind the authority's plans.

I defined public justification as the authority explaining what its plans are for and shedding light on their plausibility as solutions to societal problems. I also said that we need public justification on the part of the authority for the subjects to assess its plans. We might think that there is something suspicious or wrong about someone presenting us with plans that we need to accept at face value without the possibility of evaluating them. So, we might be tempted to add an epistemic requirement that the subjects evaluate the plans the authority is offering and their suitability as solutions to societal problems. But we need to consider that epistemic justification as an evaluation on the part of the subjects appears to pose two problems for the delegation theory. I explain them in turn.

Firstly, a general point about the epistemic requirement. An epistemic requirement that the authority publicly justifies its plans – no matter what form it takes - might seem to undermine the point of delegation. If I am delegating some tasks, I am relying on someone else to get things right on my behalf without checking that she is getting things right. In the case of political societies, we delegate social planning to the authority, and we rely on it to plan on our behalf without checking at every turn how the authority is planning. The point of delegation is to lighten our workload by offsetting tasks we do not have time for or cannot do ourselves. In political societies, we need to delegate because we cannot do everything ourselves. Delegation is, then, necessary and its point is that we outsource our plans. Engaging in an evaluation of the plans renders delegation useless because it undermines its very point: delegation is about not doing certain things yourself and allowing another to do them for you. How can we reconcile the very concept of delegation with the epistemic requirement that the authority publicly justifies its plans? This point will be clearer when I expand on the definition of delegation in the next chapter. Here, we can say that a possible solution to reconcile delegation and public justification lies within the scope of delegation itself. What is it that we are delegating, or that we have to delegate, to the authority? We might delegate the creation of the plans themselves and everything that goes into it, such as evaluating reasons against or in favour of certain courses of action, surveying possible plans to see which one is more adequate and so forth. However, we might still want to know more about what goes into the creation of authoritative plans (for the reasons mentioned earlier). If we thus define the scope of delegation, an explanation of the plans on the part of the authority would not undermine the subjects' process of delegating their planning powers to the authority.

Secondly, a point about the subjects evaluating the authority's plans. The main problem with the subjects' evaluation of the authority's plans is that it seems to clash with treating laws

and plans as exclusionary. But there is a distinction to be made between evaluating laws and plans as adequate solutions to societal problems, and the authority explaining its plans.

Evaluating plans entails reopening deliberation on the plans that the authority is giving us. This deliberation is difficult to carry out because individuals do not have the time, knowledge, and resources to survey all possible solutions to societal problems by themselves. It is also against the very concept of delegation because, as we explained, delegating is about outsourcing tasks and not having to do them yourself. When the authority explains its plans, however, it clarifies why these plans are being implemented and provides a basic explanation of their plausibility as solutions to societal problems. As we said earlier, it is reasonable that individuals know why they are asked to do certain things and what certain plans are for. When the authority explains why these plans are being implemented, it is offering a justification for its action that serves as a basis for individuals to be reassured that the first condition of legitimacy has been met. This is not the same as re-evaluating possible plans and re-opening deliberation on a certain matter. If the justification is not sufficient, or if plans later turn out to be terrible at solving a problem, deliberation can be reopened, and plans can be questioned. But individuals are not constantly evaluating and second-guessing the authority's plans. These plans should still be treated as exclusionary, and the explanation the authority gives for them merely reassures individuals that they are serving a specific purpose.

The kind of public justification that political authorities need to provide to be legitimate under the delegation theory bears some resemblance to the public reason debate (Bird 2014; Boettcher 2015; Brower 1994; Cohen 2009; Donahue 2020; Enoch 2013; Rawls 2001). For proponents of public reason, the authority's rules – moral or political - that dictate social life need to be justifiable or acceptable to its subjects. The justification of the reasons for imposing and enforcing certain rules is given by the authority through the appeal to arguments that its subjects (idealised at different levels) can accept. I do not have the space to fully flesh out the

similarities between public reasons accounts of legitimacy and the delegation theory. That is the focus of another project. However, I can say here that the delegation theory is closer to procedural accounts of public reasons, as opposed to more substantive accounts (Peter, 2007) of public reasons. Rawlsians see public reasons as reasons that attract a hypothetical consensus. Fabienne Peter (2007) argues that this consensus can be substantive or procedural. According to the substantive view, the hypothetical consensus of idealised citizens is a consensus on the political decisions made by the authority. On the procedural view, the consensus is on the procedures through which those decisions are made. Using the terminology of practical reasoning, we can interpret this as meaning that hypothetical consensus can be on the content of the authority's deliberation (substantive view) or the process of deliberation itself (procedural view). In the context of justifying democratic legitimacy, Peter defends a procedural interpretation of Rawls' idea of public reasons, according to which public reason justifies the principles that should govern the process of democratic decision-making, but it does not justify the content of public deliberation.

I am sympathetic to the procedural view of public reasons advanced by Peter as it is in line with the idea of having a practical authority that plans on our behalf. As we argued, it is in the essence of the kind of practical authorities that states are that they do not want their subjects to reopen deliberation on their decisions. Practical authorities do not want the subjects to question the content of the law or rules they put in place for coordination and compliance concerns. Furthermore, if the delegation theory is right that we ought to delegate our planning powers to authorities in political societies, it is in the essence of delegation that we outsource social planning, meaning that we do not engage in the process of deliberation on social problems and their solutions ourselves. So, if we follow these definitions of practical authority and delegation, it follows that we do not need the hypothetical consensus of idealised citizens on the content of the political decisions made by the authority. However, it is reasonable that

the subjects are reassured by the authority that its deliberation and the decisions made are tracking societal problems. If those decisions made by the authority were not tracking those problems, they would defeat the purpose of outsourcing our social planning to the state. So, it is reasonable to expect the authority to explain how its decisions are related to what we want the authority to do or the problems we want it to solve. Public reason, in this case, justifies the processes used in authoritative decision-making, rather than their content.

Before we move on to analyse the concept of delegation and how it fits into our theory of legitimate political authority, we need to say something about the constituency of public justification. To whom does the authority need to publicly justify its plans? And who is required to obey the authority? These are questions about the subjects of political authorities, which are those agents authorities need to justify their plans to and, ultimately, those who are required to obey the authority.

The first condition of legitimacy outlined in the previous section is about the relationship between the law, authorities and social planning. That condition defines the scope of social planning. Authorities, and the laws they produce in the form of plans, are directly related to social planning as authoritative plans need to track societal problems to be adequate. So, this first condition of legitimacy establishes how authorities need to approach social planning for their subjects.

The second condition of legitimacy – which requires authorities to publicly justify their plans – is about the relationship between the authority and its subjects. Unlike the first condition, this second one brings the subjects to the forefront of the question of legitimacy. If the authority is required to publicly justify its plans, there must be someone to whom these plans need to be justified. And if the authority is legitimate, it will be legitimate over someone or other. As I will explain in the next chapter, the constituency of public justification matters

for delegating planning to political authorities as well because it is important to understand who has to respond to which authority. In the case of normative delegation – according to which agents have a duty to delegate their planning powers to the authority when it comes to social planning – this becomes the question of who will be treated as having delegated to the authority.

So, who has to obey the authority? And to whom does the authority need to justify its plans? The clear-cut, paradigmatic case would be one where the world is divided into different territories and each territory has an authority with a monopoly of power and coercion over all those who reside in that specific territory. In this simplified world, no one leaves the state one was born into (and subjected to) if not for short periods (e.g., for holidays or brief work trips). In such an idealised scenario, the state exercises its power over its citizens who, as we assume, do not ever leave the state's territory for too long. The constituency and target audience of the public justification is, then, made up of those who were born in the state's territory, as well as those who occasionally pass through it as tourists.

However, we know that the world is much more complex and that what we dubbed the 'paradigmatic case' is usually accompanied by a host of complex cases. It will be impossible to identify and mention them all, but we can try to clarify a few of those cases. Firstly, states are normally authoritative over both citizens and residents, those who have citizenship and those who live and work in its territory without having that status. Citizens and residents have different rights and responsibilities but what they have in common is the fact that they need to respond to the authority of the same state and follow its plans. Secondly, some individuals will have dual nationality, where they can be citizens of two (or more in some cases) countries at the same time. Dual citizens will have to primarily respond to the authority whose territory they reside in, because, we may assume, they spend most of their time there and they need to abide by more plans in that territory. However, at the same time, they are also under the authority of the other state of which they are citizens, and they need to follow its plans regularly.

An example of this is people with dual nationality having to pay taxes in both countries they are citizens of. Another complex case is that of tourists, who primarily respond to the state they are citizens of, but who also follow the plans of the authority whose territory they are visiting. This is not an exhaustive list of complicated cases of the constituency of public justification and delegation. We can, though, try to generalise from these cases to establish some rules. More will be said about this in the next chapter but, given that agents have a duty to delegate their social planning, they will always have to respond to some authority. Conversely, the authority's plans always need to be imposed on and justified to a set of subjects (which, in our complex, modern world is bound to change at all times). So, the rule of thumb is that the authority needs to justify its plans to all those who reside in its territory, whether they are citizens of the state or not. From the subjects' perspective – which I will expand on in the next chapter – they always need to respond and delegate to the authority in whose territory they reside. It is also the subjects' responsibility to ensure they are not violating the laws of countries they briefly visit (for tourism or other purposes). In Chapter 7, which deals with delegation, I expand on the question of which authority the subjects ought to obey.

6.3 Summary

In this chapter, I argued that political authorities are legitimate when they meet two conditions:

- a) when they provide their subjects with adequate plans that help them respond to reasons;
- b) when they provide a basic public justification for their plans.

In section 6.1, I explained that plans are adequate when they are sensitive to solving societal problems and, in doing so, they need to account for and respect all the relevant moral considerations. In section 6.2, I analysed the public justification condition of legitimacy and

explained that the authority needs to explain what its plans are for and say something about their plausibility as solutions to societal problems. When political authorities meet these two conditions, they are *de jure* legitimate, whether their subjects see them as such or not. In the next chapter, I argue that the subjects' duty to delegate their planning powers to the authority is activated when the two conditions above are met.

Chapter 7 - The Delegation Theory (Part 2): A Defence of Normative Delegation

7.0 Introduction

In Chapter 6, we started unpacking the conditions of legitimacy for political authorities according to the delegation theory. We said that states are legitimate when a) they provide their subjects with adequate plans that help them respond to reasons, and when b) they accompany those plans with a basic public justification that explains what the plans are for. The chapter defined the desideratum for adequate plans and continued the argument started in Chapter 5 on the transmission of reasons from the authority to the subjects. I argued that when political authorities meet the two conditions above, they are *de jure* legitimate, whether their subjects see them as such or not.

I have contended that, when those two conditions of legitimacy are met, the subjects ought to delegate their planning powers to the authority (at least when it comes to social planning). This chapter analyses delegation and what delegating planning powers to the authority entails. Here is how I will proceed. Firstly, I define delegation, with an emphasis on delegation in the context of political societies, and briefly canvass how it has been analysed in the literature (section 7.1). I then distinguish between delegation and consent, which are two similar mechanisms that alter the normative relationship between two parties (section 7.2). In section 7.3, I introduce normative delegation, that is, the duty to delegate that is activated for subjects of political authorities under certain conditions. Normative delegation, which is based on David Estlund's notion of normative consent, is the opposite of voluntarily delegating tasks to someone, and section 7.3 explores its workings and implications. Finally, in section 7.4, I address some objections that are frequently raised against normative consent to see whether

normative delegation, which bears similarities with the notion of consent in general, falls prey to the same criticism.

7.1 Delegation

In this section, I analyse and define the phenomenon of delegating social planning to political authorities. I start by giving a brief overview of the literature on delegation and proceed to provide a general definition of it and a more specific one that applies to delegating tasks to the state.

The notion of delegation has mostly been analysed – to my knowledge – in three contexts: economics, philosophy of law, and a strand of political philosophy, democratic theory.

In economics, the process of delegating tasks to others has been studied in organisational contexts, where managers (or superiors in general) delegate tasks to other staff members or their subordinates. In this context, delegation is sometimes framed in terms of the differences between centralised decision-making, where the power and control reside with upper-level managers, and the outsourcing of tasks to lower levels of the organisation's ranks. With delegation comes a loss of power at the higher levels but (in some cases) better use of time, resources, and skills. Notable in this area is a study by Angst and Borowiechi (2013) that defines delegation as a decision rights transfer whereby the agent who gets assigned tasks via delegation has now the right to choose how to carry out projects or duties.⁵⁶ Delegation and its

⁵⁶ See Angst and Borowiechi (2013) for a focus on the effects of delegation on the agent's motivation when she gets assigned tasks. A similar previous study is the one by Aghion and Tirole (1997) who describe the principal-agent game Angst and Borowiechi based their laboratory study on. In the principal-agent model, the principal is the one who delegates tasks, and the agent is the person tasks are assigned to. Aghion and Tirole see delegation as a tool to allocate decision-making rights and authority within organisations and their model is used in other studies. See also Swank and Visser (2007) for another economics study on delegation.

psychological effects are studied in economics to understand the power dynamics and functioning of organisations with various levels of hierarchy.

In philosophy of law, the phenomenon of delegation is analysed mainly in relation to states and lawmakers delegating powers and responsibilities to supranational organisations or legal officials. In this context, delegation is understood as a transfer of power to execute tasks from the states to supranational organisations or legal officials. When it comes to international criminal tribunals and their power or authority to judge crimes, for instance, we might ask where that authority comes from. One possible ground for international criminal tribunals' power to judge crimes and punish wrongdoers can be the transfer of that specific power from states (Wallerstein, 2015). This is an example of how states can use delegation as a tool to transfer powers that belong to them to other organisations.⁵⁷ Delegation as a transfer of powers can also occur from legal official to legal official (Asgeirssons, 2013). This is a useful tool which legal officials can use to maximise their time and resources. For some (Asgeirssons, 2013), delegation is justified even when the delegates are not better suited to solve the problems in question. For others (Sorensen, 2001) delegating powers to officials can be justified only if these officials are better suited to solve or answer certain cases. This legal literature on delegation is important because it allows us to better understand the function of delegation and what agents can achieve with it. We can learn from these models and adapt them to the case where the subjects of the state delegate some of their powers to the state itself.

In political theory, as De Smet, Peeters and Sterckx (2016) argue, the delegation of power and responsibilities from the citizens to the state is a theme that has been analysed only implicitly. De Smet *et. al.* are interested in the citizens delegating the task of tackling climate change to the state, which is arguably better able to solve this complex issue that is difficult to

⁵⁷ Cfr. Luban (2009) for an opposing view about the inherent problems with states transferring their power to punish to other organisations.

tackle at the individual level. They found a more explicit formulation of what they call the *delegated authority model* in the case of climate change action in Gardiner (2011). Gardiner refers to a long tradition in political theory where citizens delegate powers and responsibilities to states, and that delegation makes those states legitimate. De Smet *et. al.* (2016) then proceed to develop an account of responsibility according to which delegating tasks to others (states, in our case) does not relieve the delegating agents from responsibility related to action or inaction on the issues in question.

We can trace back the (likely) origins of what De Smet calls *delegated authority model* to Rousseau. Rousseau (1762) argues that the citizens – ‘the people’, as he refers to them – are the sovereign with political power. The sovereign delegates this political power, which resides in their hands, to legal officials (the government) that can exercise that power on the citizens’ behalf. For Rousseau, the citizens delegate executive power to legal officials. To refer to Aghion’s and Tirole’s (1997) principal-agent model, the sovereign in Rousseau’s account is the principal, the one who delegates, and the government is the agent, the one who acts on behalf of the principal. Political power is always in the hands of the sovereign for Rousseau, but the fact that it can be delegated to legal officials allows a specific form of government – representative democracy – where people can retain their sovereign powers even when they do not directly participate in political life and others act on their behalf (Downey, 2021).⁵⁸ Delegated sovereign power is useful to reach common goals and carry out complex tasks that require an extensive organisation that is not possible to implement at the individual level.

Rousseau’s account of legitimate authority paved the way for modern studies of democracy and democratic power. In democratic theory, different aspects of delegation are

⁵⁸ We need to note that Rousseau was not sympathetic to representative democracy, so he would not endorse this particular use of delegated powers. However, some (Downey, 2021) have emphasised the influence Rousseau had on the representative democracy debate.

analysed, one of them concerning legislators transferring some of their policy-making powers to administrative agencies, such as the central bank in the case of monetary power (Downey, 2021). A host of studies analyses the limits of delegation and good delegation practices. Richardson (2002) uses public reasoning on the part of the state as a way to constrain and regulate delegated power. Christiano (2005) also appeals to public reasoning as a way to regulate delegation and also adds that legislators need to retain complete control over the policymaking ends when they delegate policymaking powers to other institutions.

There is a lot to learn from all these approaches to delegation in economics, legal theory, democratic theory and political philosophy. I am specifically interested in the kind of delegation that goes from the people to the state, where the subjects of a state delegate their social-planning powers, powers to organise and coordinate their collective actions, to political authorities. Despite this being, at its core, a democratic idea, I use this type of delegation as a ground for political authority in general, whether that authority is organised in a form of democratic government or not. The delegation theory does not present, in this respect, a novel or revolutionary ground of political legitimacy. However, it gives an explanation of what delegation is in terms of the theory of action – something that has not been done before, to my knowledge – and how it differs from other mechanisms that alter our normative relationship with others (i.e., consent).

I provide now a broad and general definition of delegation. Delegation can be defined as an act⁵⁹ through which one or more agents (or institutions) authorise other agents (or institutions) to perform tasks on their behalf. The agent who is outsourcing tasks is the mandator, whereas the person who assumes the responsibility of acting on the mandator's behalf is the mandatory. By requesting others to do something on their behalf, mandators

⁵⁹ In small-scale settings, this is likely an express, concrete act (as opposed to a hypothetical act).

transfer their power to do certain things to the mandatories, who are then authorised to act for them in different ways, such as signing or speaking in their place, and performing tasks on their behalf (Bourdieu, 1985).

Embedded in delegation are the concepts of authorisation and representation. By delegating tasks, mandators authorise mandatories to do certain things for them, thus vesting them of the power and right to act in ways that were unavailable to them before. If I delegate the organisation of a workshop to one of my interns, I authorise her to use the data and materials of the organisation – which I have the power and right to use – to do something on my behalf. The intern could not rightfully make use of that data and materials without my authorisation. When delegation involves the performance of specific acts, such as signing a document or speaking on the mandator's behalf in public, the performance of those acts on the part of the mandatory involves representing the person who authorised them to perform those acts. If I am unable to vote in person due to an illness, I can authorise my friend Maggie to vote for me by telling her what candidate or party I want to vote for and by signing the relevant paperwork. In crossing the name of my preferred candidate on the ballot paper, Maggie is representing me whilst acting in the way she was delegated to act.

As Arthur Lupia (2001) points out, the main benefit of delegation is that it allows individuals to maximise the number of tasks they can accomplish by overcoming personal shortcomings, such as limited time, talent, and energy. Delegating tasks to others, then, allows individuals to do more by tapping into other agents' time, talent and energy (Lupia, 2001, p. 3375). Delegation takes place between two parties but can take different forms depending on who is delegating to whom. Single individuals, groups or institutions can delegate to other individuals or institutions. Even governments, as we discussed before, can delegate tasks and responsibilities to supranational organisations, such as the European Union, or international organisations, such as the World Trade Organisation.

The type of delegation we are interested in is the one that goes from individuals to states, that is, from individuals to political organisations. Building from the general definition of delegation, we can now sketch a definition of political delegation, which, as I will argue, partially grounds legitimate political authority. In political delegation, the citizens and residents of the state (therein, the subjects) are the mandators – those who delegate – and the state is the mandatory – the organisation to which tasks are delegated. Political delegation is, then, an act through which the subjects authorise⁶⁰ the state to perform tasks on their behalf. Political delegation, at least in its voluntary form, can be an express or tacit concrete act that the subjects perform to outsource tasks to the authority. I talked about ways in which individuals can signal their tacit consent to the state in Chapter 2, so I will be brief here. Individuals can delegate in an express way when they have the chance to vote or personally request that the authority does things for them verbally or in writing.

It is worth considering that in large-scale setting it is costly to continuously provide mechanisms for group members to express their delegation to the state. For this reason, a form of tacit concrete act of delegation might be preferred, such as residing in a territory, remaining silent or simply showing obedience to the authority.⁶¹ This means that individuals can signal that they are delegating their planning powers to the authority by obeying it and following its plans. When citizens and residents go along with authoritative plans, and the authority makes it clear that their cooperation with the law amounts to tacit delegation, the authority can operate with the assumption that it has the tacit delegation of its subjects. This will inevitably lead to making some mistakes, as citizens who have not tacitly delegated might be treated as if they did. Tacit consent can be overridden by an exercise of active dissent. If the subjects wish to oppose to the authority or make it clear that they did not tacitly delegate, they need to express

⁶⁰ As we will see later, there is a voluntary component to delegation when individuals spontaneously authorise others to do things on their behalf.

⁶¹ But bear in mind that, ultimately, I argue for normative delegation rather than voluntary delegation.

their dissent.⁶² I will elaborate on whether it is wrong to assume the subjects have tacitly delegated and act accordingly when I analyse normative delegation. Here I need to highlight that both those who tacitly delegate and those who do not have political obligations towards the state (obligations to obey its laws). Those who delegate have obligations to obey the law that stem from their act of delegation. Those who do not delegated are still obligated to obey the law and punished if they do not because their disobedience can hinder those who delegated. If there are many free riders who benefits from the state without contributing (without delegating), this can have negative effects on social order and damage the portion of the population that delegated. I will talk more about this below when I discuss and defend normative delegation.

Similarly to general delegation, the subjects do not have the time, resources or skills to perform those tasks themselves. Members of political societies need their actions to be coordinated with the actions of others in a way that is conducive to harmonious co-existence in the state's territory. The social planning required to organise the actions of people in political societies is so wide-scale and time-consuming that individuals cannot tackle it by themselves unless they cooperate and act with others to work towards it. Cooperation with many people might not be possible to execute by agents without an external organisation (or institutional body) that regulates them. Even if it were possible to execute, that coordination will require resources and time that individuals do not possess. To overcome this shortcoming, they can delegate the task of social planning (i.e., coordinating their conduct within the context of living in the political society) to the state, which has the time, skills and resources to execute large-scale social planning. Through the act of delegating social planning to the state, the subjects

⁶² I will not elaborate on which ways would be best to express dissent or how the subjects can effectively dissent in political societies. This will be the object of a different research project.

transfer their planning powers to the state, which is then authorised to create and enforce plans the subjects have to follow.

The authorisation embedded in the act of delegation, then, gives the state the power and right to plan on behalf of the subjects and gives the subjects a duty to obey the state plans. In cases where, after the subject's delegation of their planning powers, the state plans in a way that follows the subjects' wishes or reflects how the subjects themselves would plan, the state discharges the delegated tasks whilst also representing the subjects. A form of democratic government, for instance, would be built on this kind of representative delegation.

7.2 Delegation vs. Consent

We defined political delegation as a form of authorisation whereby the subjects give the state the power and right to do something on their behalf. The primary task the state carries out for its subject is social planning, that is, the creation of plans to regulate social life. Delegation seems to share this authorisation component with consent. Consent is generally taken to be a mental state or act capable of transforming the normative landscape between two parties (Kleinig, 2009). Something that would not be permissible in the absence of consent becomes permissible when consent is given (Kleinig, 2009, p. 4), as it happens when someone gives consent to sexual intercourse which otherwise would be rape. Delegation functions in a similar way. When we authorise someone to act on our behalf, we enable that person to do something she could not do absent that authorisation. For instance, through delegation, my accountant can organise and file my tax report, which she could not rightfully do before I delegated that task to her. Delegation and consent share the ability to alter a normative landscape and make actions right that would otherwise be wrong. Moreover, much like delegation, consent is often described as a way for someone to authorise someone else to do something, such as when I

consent to my brother using my car for the day, where my consent is an authorisation for my brother to borrow my car (Kleinig, 2009, p.8).

Just as delegation takes different forms, as we saw earlier, consent too has a broad scope, as we can consent to different types of things and actions. We can consent to someone doing something to or with our bodies (e.g., consent to medical procedures or sexual intercourse), to someone using our property (e.g., to my brother borrowing my car), or we can consent to do things for someone else (e.g. when someone consents to give a speech at a wedding). John Kleinig (2009) notably argued that the two main ways consent works are by being a gate that we can open to give someone else access to us⁶³ (e.g., consent to medical operations, sexual intercourse, use of our property, dissemination of our personal information); or by forming a normative rope that binds us to other agents (e.g., consent to marriage, to doing something for others) (Kleinig, 2009, p. 4). Delegation can also function as a normative gate or rope. When we authorise another agent or institution to do something on our behalf, this authorisation creates a normative rope – to use Kleinig’s terminology – that binds others to us. For instance, if we delegate the organisation of our birthday party to a friend, we bind our friend to us by imposing on her the obligation to organise the party (provided she accepted the terms of delegation). In this way, delegation works as a normative rope by binding the person or group that tasks are delegated to the ones who delegated. At the same time, this binding act also gives our friend the power and right to do the task we assigned her. We thus empower her to act on our behalf and take steps she might not have had the authorisation to take before. This is how, by delegating, we can open a normative gate to allow people to do or say things on our behalf, a power they did not have before the act of delegation.

⁶³ Note that the access that consent gives others to us, or how it binds us to others would not be permissible before the act of consent (Kleinig, 2009, pp. 3-6).

Given that both delegation and consent are forms of authorisation, we might be wondering what the difference between them is, if there even is a difference. I argue that their shared normative powers (to bind other people to us and open gates for them to do things to us or on our behalf) have a different target. Consent seems to be more individualistic in nature, as it works to protect our autonomy and bodily integrity from those who want to violate or breach it (Beauchamp, 2009).⁶⁴ The target of consent is a more individualised dimension where there is something wrong with accessing someone else's body or property unless we obtain the consent of the other person. In this respect, consent is a normative device that protects us as individuals. We can close the gate that leads others to our bodies and our possessions if we do not wish something to be done to us or with our things. Delegation seems to track a different dimension that is inherently social and collective.⁶⁵ By consenting, we give people the green light to do something to us or with our things. Delegation, as another form of authorisation, also allows us to give other agents the green light to do something. However, delegation is not just about authorising people to do something: it is also about why we are giving the green light and what we can achieve in doing so. By delegating tasks to other people, we are outsourcing tasks that we could either do ourselves but do not have the time to complete, or tasks that we cannot do as individuals for lack of skills or resources. When we delegate to other people, we transfer our agential powers to someone else, that is, the power we have to act in certain ways gets transferred to another agent or group that is authorised to perform those actions on our behalf.

Our agential powers, which we can transfer to another via delegation, exist in a collective and social dimension that differs vastly from the individualised bodily space that consent aims to protect. We can understand why if we look at the consequences of not delegating versus not

⁶⁴ For a more detailed analysis of bodily integrity see Blackman (2010); Shildrick (2010) and Viens (2014).

⁶⁵ Although note that the interpretation of consent as a normative mechanism that binds us to others (Kleining, 2009, p. 4) can also be said to have this collective flavour.

consenting to something. Except for post-mortem organ procurement, when we deny someone access to our body we are not damaging or negatively affecting other agents. If I refuse to permit my doctor to operate on me to remove a pesky mole, I might damage myself and my health should it turn out that that mole was a malign tumour. But I am not necessarily damaging any other agent by refusing that the doctor does something to my body.⁶⁶ When it comes to refusing to transfer my agential powers to someone else, however, the implications of my action or inaction on other people are more obvious. When we refuse to delegate our social planning powers to the state, for instance, our doing so might negatively affect other agents if our delegation is required for the state to do something of public interest. If enough people refuse to delegate to the state, and the state does not perform an essential function as a result of that refusal, not delegating can impose harm on their fellow group members.

Our agential powers are pointed at a collective space that we share with other agents in our society. We cannot exercise our agency in ways that would exclude others. The purpose of delegation is to allow us to do what we like, but it involves the management of agential powers that are less individualistic in nature than the dimension of bodily integrity that consent protects. So, authorising other agents to do things to us or on our behalf, which is the common ground between delegation and consent, has a different target when we consent to something than when we delegate our agential powers to someone else. When we give our consent, that form of authorisation has an individualistic dimension that protects our bodily integrity and allows others to temporarily access us or our possessions. When we delegate our agential powers to others, that authorisation is directed at the social and collective dimension where our agential powers interact with those of others. This is the difference between delegation and

⁶⁶ Except, though, opportunity costs: I might require more expensive medical intervention later on.

consent.

7.3 Voluntary vs. Normative Delegation

In defining and describing delegation, I referred to it throughout as something voluntary, an authorisation we freely give to others so that they can perform tasks on our behalf. In this section, I consider normative delegation, which, as I defined it, is the duty to delegate that activates for subjects of political authorities under certain conditions. Normative delegation is, in this respect, the opposite of voluntary delegation. When we voluntarily delegate, we choose to authorise someone to do things on our behalf. This authorisation gives the delegate the power to do things that they otherwise would not have been able to do in our name. When it comes to normative delegation, we do not get to freely authorise people to do things for us: we have a duty to delegate and the person or institutions who need our authorisation to do things would treat us as if we indeed delegated. In this section, I argue that normative delegation is a better candidate to legitimise political authority than voluntary delegation.

The definition of (political) delegation provided earlier, and how we approached it so far, presupposes a voluntary transfer of powers, rights and responsibilities from the mandator(s) to the mandatory. In cases where the mandatory is a political practical authority (i.e., a state or political organisation) this voluntary act (transfer) results in the mandators having to conform to the mandatory's requests or directives, and the mandatory enforcing those directives to maximise compliance with them.⁶⁷ With voluntary delegation, once the mandator transfers

⁶⁷ Note that delegation works differently in cases where theoretical authorities – as opposed to practical authorities – are involved. Theoretical authorities are experts in their fields who can advise others on how to behave and can say what is or is not the case in their field. Theoretical authorities give others reasons to believe what they say or advise but they do not have the power to make others act as they wish (Ehrenberg, 2011). Here, mandators can delegate certain tasks to experts who have a better chance of completing them successfully. Those experts can deliver a finished product that mandators might use or benefit from, but experts do not and cannot force mandators to act in certain ways. If I seek the services of a nutritionist to improve my diet, I delegate the creation of a nutrition plan to an expert. The expert will provide me with that service, and it is up to me whether to conform or not to her advice. It might be foolish for me to ignore it since it would be a waste of my money and the expert's time, as

powers, rights or responsibilities, they are treated in a certain way, that is, they are presented with directives that they need to obey, and compliance with those directives is not only expected but also enforced.

The problem with voluntary delegation which, in theories of political authority, would underpin the legitimacy of states, is that it requires some sort of voluntary act or commitment on the part of the citizens for delegation to occur and the authority to be legitimate. And we learned from consent-based accounts of political legitimacy how hard it is to obtain the consent of the majority in a large-scale group and, in some cases, to ascertain that consent was given (when, for instance, people tacitly consent to something) (Simmons, 1976; 1998).⁶⁸ Being very similar to consent and the way it changes agents' normative landscape, delegation – the voluntary kind – suffers from the same issues consent theories have tried to solve for years, such as what kind of acts count as delegating to the authority, how to establish whether citizens have delegated if the action they need to perform could be interpreted in different ways, and so forth. Moreover, sometimes people simply do not show up to give their consent or delegate to the authority. In the 2019 General Election, only 67.3% of the UK population voted, leaving out 32.7%. How is the state to treat that 32.7%? Should that portion of the population be treated as having delegated (or consented) to the authority, thus giving them a duty to obey the authority? Or should that 32.7% be considered dissenters? And, if so, what changes in their treatment compared to the 67.3% who delegated? These are only a handful of the many questions that voluntary theories of political legitimacy pose. We certainly value citizens' opinions and for them to be able to express them to exercise their autonomy (among other

well as detrimental to my health. However, upon being presented with the nutrition plan, I am under no obligation to follow it. So, in cases of delegation where theoretical authorities are involved, there is still a transfer of responsibilities from some agents to other agents, but there is no obligation on the part of the mandators to do as they are told. Obligations might arise for the mandatories to deliver their services after they enter an agreement with the mandators.

⁶⁸ Recently, there has been a renewed interest in the possibility of tacit consent, with authors such as Cibik (forth.); Furner (2010) and Puryear (2021) trying to defend it or somehow re-evaluate the theory.

things), and we spent the first part of this dissertation figuring out if and how citizens maintain their autonomy in cases of hierarchical shared activity.

But sometimes, the problems needed to be solved are too pressing for the authority to convince its subjects that they need to delegate and wait for them to do so. The urgent task approach, as David Estlund (2008) calls it, emphasises this point. According to it, some tasks are so pressing and important that people have an obligation to obey the authority, or whoever steps up to be an authority in an urgent situation, that can solve or reduce the pressing problems (Estlund, 2008, p. 147). Some cases where there is an urgent problem to solve might look like *Beach Rescue*, the case we analysed in Chapter 5, where a timely, concerted effort is required to prevent something bad from happening immediately (to save a life, in that case). But other cases might involve something of crucial importance that needs to happen over a prolonged period. As we argued, continuously coordinating the conduct of subjects in political society is of paramount importance for the functioning of the society itself and the resolution of societal problems. The kind of urgency associated with social coordination has to do with the importance of this task, and the tremendous efforts that need to be put into sustaining it overtime.

The urgency and importance of delegation, coupled with its role in making political authority legitimate, raise some interesting questions that I would like to explore here. Could there be cases where agents are treated as if they delegated to someone despite the absence of actual delegation? And do subjects in political societies have a duty to delegate their social planning powers, as we called them, to the authority?

In what follows, I argue for what I call ‘normative delegation’, a non-voluntary approach to delegation that is modelled on David Estlund’s (2008) account of normative consent and its adaptation to the context of opt-out systems of organ procurement by Ben Saunders (2009).

The main idea of normative (political) delegation is that subjects of practical authorities in political societies are treated as if they delegated their planning powers to the state even when delegation itself did not occur on their part. If the subjects have a duty to delegate social planning to the authority – which they do have, as we will see – the authority is justified in treating them that way even though they did not engage in any act of delegation. The duty to delegate planning powers to the authority – an idea that is connected to urgent task theories – stems from and is related to a more general duty to coordinate actions with those of others in social contexts. Let us unpack these ideas further.

The starting point of the argument is that the actual act of delegating something to the authority is not required when it would be morally wrong for agents not to delegate the thing in question to the authority. When it is morally wrong for people not to delegate in certain circumstances, they have a duty to delegate in those circumstances.

This is based on Estlund's analysis of symmetrical cases in which both consent and non-consent can be nullified. As a normative rope that binds or a normative gate that gives access, consent can give rise to obligations between two parties. However, there are cases in which even actual, explicit consent can fail to produce obligations, such as when someone consents to become someone else's slave. Even if one gives her voluntary, fully informed, non-coercive consent to becoming a slave, for instance, that act of consent is invalid and will fail to generate an obligation to enslave oneself to others because slavery is morally wrong. And we think that we cannot or should not consent to immoral things.

Estlund (2008) wonders whether, similarly to invalid instances of consent, non-consent can also be morally wrong and nullified under the right circumstances. To show that non-consent can be nullified sometimes, he analyses a case that is similar to *Beach Rescue*. Estlund's case involves a passenger aeroplane that has crashed and in which a flight attendant

steps up to give commands to people to save as many passengers as possible. There is a good chance of saving many passengers through a very well-coordinated effort and we have someone who organises this rescue procedure. If the flight attendant is not giving immoral commands, and if the efforts of many are required to save people, Estlund argues that it would be immoral not to obey her commands. It would be immoral for someone able to help not to agree to go along with the attendant's commands in this case. If people refuse to consent to obey the flight attendant, their non-consent is invalid because it is immoral for them not to consent to obey in this case, since it is immoral not to help people in need if one is in a position to do so. When non-consent is invalid, people's permission or authorisation is no longer sought or required to do something or act in certain ways. Agents who did not consent – where the non-consenting is invalid – are treated as if they had consented. In cases where authorities are involved, authorities will treat non-consenters as if they had given their consent, and agents will be under an obligation to obey the authority. This is, for Estlund, the basis of normative consent.

I believe we can apply Estlund's argument for normative consent to delegation. As we already mentioned, delegation and consent share many similarities: they can both generate rights and obligations in changing the normative landscape between parties, and they both function as a normative rope, with delegation being a rope that binds others to us. It is safe to say that, in some cases, delegation is null and fails to generate the standard obligations that it would otherwise generate. When Denzel hires an assassin to get rid of Geoff, his business rival, Denzel delegates the task of killing Geoff to someone else who can do the deed on his behalf. A normal case of delegation would put mandatories under an obligation to carry out the task that was delegated to them while giving mandators the right to demand that the job is done on their terms. Here, however, we do not think that Denzel's act of delegating Geoff's killing to an assassin puts the assassin under an obligation to kill Geoff, nor do we think that it gives Denzel any right to demand that the killing gets done. A killing of this nature is an immoral act.

Just as no one should consent to do (or be subjected to) immoral things, such as consenting to kill someone, agents should not delegate immoral things to others. Consent and delegation function similarly in this respect. So, if consent to doing immoral acts is void, delegating immoral actions to others should also be void. In the case of consenting to do immoral things, the normative rope fails to bind us to others and to produce obligations for us. In the case of delegating immoral tasks to others, the normative rope fails to bind others to us and give rise to adequate rights and obligations.

Estlund found a case of symmetry between invalid cases of consent and invalid cases of non-consent. Now, is there a similar symmetry between cases of null delegation and null non-delegation? If the answer is yes, just as there are cases where the act of delegating certain things is void and fails to produce the relevant rights and duties, there will also be cases where failures to delegate are nullified under the right circumstances.

Consider this case⁶⁹.

MAZE: Thomas and other teenage boys (the Gladers) found themselves in the Glade, a large green area surrounded by tall stone walls and an intricate maze. They have no memory of how they ended up there. They just know that one by one the boys have been placed there via an underground elevator by a mysterious organisation called WCKD. Thomas, the last to be placed in the Glade, starts to remember that he worked for WCKD, which created the maze where the boys have been the test subjects for an experiment to find the cure to a mortal virus. Horrified at the fact that WCKD kept them prisoners for years in the name of science, the Gladers need to find a way to escape. To do so, they must work together. The Gladers know that for his (regained) internal knowledge of

⁶⁹ Loosely based on the 2014 20th Century Fox movie *The Maze Runner*.

WCKD, his bravery and moral strength, Thomas is in a much better position at guiding them through the maze than anyone else. Following Thomas gives them a better chance at escaping than trying by themselves. Some of the boys delegate the planning of their escape to Thomas, while some refuse to delegate, as they do not trust him. Various attempts at escaping fail because of these internal tensions, with the result that people start dying in the maze. Thomas then imposes his coordination efforts, which lead to the survivors escaping the maze.

In this example, there is an important problem (escaping the maze and evil WCKD) which requires concerted efforts from all people involved to be solved. We assumed that only one person in the group (Thomas) can solve this problem by directing others in certain ways and that there is a much better chance of succeeding guided by him than there is without his guidance. Given this person's unique ability to solve an important problem, and that the directed efforts of everyone are required to succeed, it would be immoral for the others not to delegate planning to this person. In *Maze*, when some of the Gladers refuse to delegate the escape plan to Thomas, people start dying as a result. It is important to escape to survive and they can only succeed in doing so by following Thomas. Staying in the maze puts the Gladers and others in danger, as WCKD has failed to find a cure for many years and caused many deaths. It is immoral for people not to delegate in this case, as that is the only way to save lives. If people refuse to delegate, their non-delegation is invalid, because it is immoral not to do anything when a bad outcome can be prevented. When non-delegation is void, people can be treated as if they had delegated, with relevant rights and duties arising despite the absence of actual delegation.

Delegating our (social) planning powers to a political authority is much like the Gladers delegating their escape plans to skilful Thomas. In political societies, we have pressing

problems to solve, such as creating and maintaining social order, upholding social justice, compensating those who have been wronged, keeping people safe and so forth. Solving these problems requires the concerted efforts of many, and people will likely need guidance in coordination. Someone – the state as an institutional body, in this case – is better suited to solve these problems than anyone else. This is because the state has the time, personnel, expertise to draw on, and resources that are required to coordinate the behaviour of many in their territories. As Raz taught us, there is a much better chance of solving societal problems by abiding by the state's laws than there is without the state's help and guidance. If the societal problems are important to solve, and if the state has a much better chance of solving them by directing the conduct of others, it would be immoral not to delegate social planning to the state we live in. Given that abiding by the state gives us a much better chance at solving societal problems and coordinating our actions in relevant ways, it is immoral for us not to delegate social planning to the state. If agents refuse to delegate in such cases, their non-delegation is invalid, because it is immoral not to act in relevant ways necessary to prevent bad outcomes. When non-delegation is void in this sense, people can be treated as if they had delegated. This means that, in political societies, the subjects will have a duty to obey the authority, and the authority will have a right to demand compliance from them, despite the absence of actual delegation.

We found that what works in the case of consent applies to delegation. When it comes to consent, when it is morally wrong for people not to consent to something, people have a duty to consent in those circumstances. With delegation, when it is morally wrong for people not to delegate tasks to someone else in certain circumstances (most likely an authority), they have a duty to delegate (to the authority) in those circumstances. But why is there a duty to delegate social planning to the authority? And where does this duty come from?

I argue that the duty to delegate social planning powers to the authority stems from the importance of some tasks and the necessity of external coordination to achieve them. Urgent

task theories⁷⁰ stress the importance of certain tasks and how that affects how we (should) behave. They argue that some tasks are of the utmost moral importance and that we have a natural moral duty to obey the commands of the authority (which might be an established one or one that arises on the fly) that is in a much better position to achieve those tasks with our obedience than we would without abiding by it (Estlund, 2008, p. 132). Here, the fact that some urgent, and morally important, tasks can only be achieved with the help of the authority does the work in producing an obligation to delegate those tasks to the authority. Some urgent, moral tasks can be resolved individually, without anyone's help. If we take Peter Singer's (1972) example of the child drowning in a shallow pond, saving the drowning child is a task of the utmost moral importance. If I am a capable, strong adult, and if saving the child does not require sacrificing anything of comparable moral importance, it should be easy for me to pull the child out of the pond by myself. It is not morally required that I delegate this task to an authority unless ten children are drowning in the pond and five bystanders whose actions need to be coordinated to save them, which is a task that would require an external authority to be achieved. In this and similar cases, delegating the task to someone else might even be detrimental, as the time that is required for me to find help is precious time wasted that could be used to save the child. So, for the duty to delegate to arise for us when there are urgent tasks to be solved, it is necessary that the task can only be solved with external help.

Solving societal problems is without a doubt a task of the utmost moral importance, as we already explained. Since the state is in a much better position at solving societal problems than individuals, and since solving societal problems is urgent and morally important, individuals ought to delegate this task to the state.

⁷⁰ Some authors who have argued for some sort of version of the urgent task theory are Kant (1979); Anscombe (1979) and Wellman (1996).

So, the importance of some tasks and the necessity of external help to achieve them give rise to an individual duty to delegate those tasks to an authority. What is particular about societal problems, though, is that they require extensive coordination of many people's actions to be solved. Problems that involve a large number of individuals cannot be solved without this extensive social coordination. We might say, then, that when there are urgent moral problems to solve, and when it is not possible to solve them without extensive coordination (and cooperation), individuals have a duty to coordinate their actions with that of others to solve these problems. If the only way to solve problems is by coordinating, and if problems are morally significant, then we have a duty to coordinate with others to solve those problems.

Societal problems require extensive social coordination to be solved and can only be solved by an institutional body with the relevant resources, not by individuals alone. In this case, the institutional body would oversee the coordination of many people's behaviour to solve those problems. Concerning social planning, then, individuals have a duty to coordinate their actions with that of others and a duty to delegate social planning (and coordination) to the authority. Here, the duty to delegate social planning to the authority derives from a duty to coordinate your actions with that of others. The duty to coordinate our actions with others, then, generated a duty to delegate that coordination to the authority.

We said that the duty to delegate to the authority, at least in the case of solving societal problems, derives from a more general duty to coordinate your actions with those of others. But what is the basis of this general duty to coordinate your behaviour with that of others? Where does this duty come from? To understand it, we might want to turn to Locke's description of the state of nature. In Locke's state of nature, individuals – despite not being subjected to any government – still have mutual obligations towards one another, especially when it comes to not hurting others and maintaining peace. According to the law of reason, which individuals follow in the state of nature, everyone is equal and independent and has a

right to life, liberty, and property. In order not to violate those rights, people need to avoid harming others in their lives, freedom, and possessions. They have, then, an obligation not to harm other people. Coordination is the best tool to prevent harming others, in that it allows people not to get in each other's ways and to live harmoniously. Harming others is morally wrong and people have a duty not to harm others; coordination is the best way to avoid harming others so, people have a duty to coordinate their actions with those of others (to avoid harming each other).

As Simmons (1993) argues, individuals might be able to authorise third parties to settle disputes for them – and organise coordination for them - and remain in the state of nature. This means people in the state of nature might not need political authority.

Now, the state of nature for Locke can be relatively peaceful if people avoid hindering each other's rights (by coordinating their actions, we might add). However, things can nonetheless go wrong in this mostly peaceful state, and people can precipitate into a state of war if they stray from the law of nature that orders them to respect each other's rights. That is why having a political authority that can judge all disputes and ensure people respect the law of reason is useful, and individuals can agree on that and consent to it via a social compact for Locke. Regardless of whether we want to espouse the social compact view, what matters is that political authorities rely on coordination too to make sure people do not harm others. Authorities impose coordination rules on people when they are in power over them. So, whether individuals organise that coordination themselves, or are coordinated by an external political authority, they always have a duty to coordinate their actions with that of others.⁷¹

⁷¹ It is worth noting that Locke does not dispense with consent. He does not completely appeal to the value of the interests that will be protected if there is political coordination. Locke is dependent on a theory of tacit consent. It would be unfair to say that this explanation of where the duty to delegate comes from is Lockean in spirit but it does not completely reflect Locke's account of political authority.

One might not be satisfied with the Lockean picture and might want to find a different basis for the general duty to coordinate. One who espoused an anti-Lockean basis for this duty would argue that, despite individuals having certain rights in the state of nature (life, freedom and property rights), most social aspects of our rights are authority-dependent. This is evident in the case of property rights. It is the political authority that decides the extent, the boundaries, the limits and the regulations of our property rights. Without the authority regulating those rights, individuals would not be able to properly exercise them, in such a way that everyone's property rights can be similarly protected. The authority can impose social coordination on us to facilitate the protection of these rights and maintain the social order. However, insofar as we refuse to delegate our planning powers to the authority, we would not be able to benefit from those aspects of our rights that are protected and regulated by the authority. In other words, individuals cannot be completely self-sufficient without delegating some tasks to the authority. Delegating our coordination plans to the authority, then, is not only socially beneficial to avoid harming others and protecting our rights. It is also necessary if we want to build and live in recognisable, sophisticated societies.

I hope to have shown that, both on a Lockean and an anti-Lockean approach, individuals have a duty to coordinate their actions with others. It is very plausible that large-scale social coordination requires the intervention of a highly organised institutional body to be upheld. That is why we also have a duty to delegate our social planning powers to political authorities.

The importance of coordinating our actions with that of others may, then, justify political authorities creating and imposing social plans on us. In the first part of the thesis, I argued that acting with others and abiding by plans created by someone else is compatible with acting autonomously. If I succeeded in showing that abiding by authoritative plans is compatible with our personal autonomy, delegating our social planning powers to the authority, then, should not be too onerous on us or alien to our nature as social beings.

7.3.1 The two conditions of legitimacy and delegation

Delegating social planning powers to the authority is, of course, not enough for authorities to be legitimate. And we previously laid out two conditions of legitimacy. Now we are going to explain the connections between delegation and the two conditions of legitimacy.

Previously, we said that political authorities are legitimate when they provide their subjects with adequate plans that help them respond to reasons (ADEQUATE PLANS); and when they give a basic public justification for their plans (PUBLIC JUSTIFICATION). ADEQUATE PLANS and PUBLIC JUSTIFICATION are connected because authorities need to motivate their choices (at least minimally) when it comes to plans for the subjects to be able to ascertain that those plans help them respond to reasons. Normative delegation is tied to ADEQUATE PLANS because individuals do not have a duty to delegate when the authoritative plans are inadequate. ADEQUATE PLANS and the duty to delegate go together in two directions. When ADEQUATE PLANS is satisfied, agents have a duty to delegate their planning powers to the authority. When ADEQUATE PLANS does not hold – because authoritative plans are inadequate – there is no duty to delegate on the parts of the subjects. The duty to delegate, then, follows directly from ADEQUATE PLANS. If the social plans created by the authority are unjust, immoral, or are not properly tracking social coordination, the subjects' duty to delegate will not 'activate' and they will not have or be treated as if they have obligations to the authority. Individuals do not have a duty to delegate when that would entail following improper plans. These are the connections between the conditions of legitimacy and the subjects' duty to delegate.

It is now time to analyse an issue that we touched on at the beginning of this chapter. The two conditions of legitimacy I offered at the beginning of Chapter 6 (section 6.0) have different scopes. (a) is about the relationship between the law, political authorities, and social planning.

It establishes how authorities should plan for their subjects and the functions of authoritative plans. (b) is about the relationship between the authority and its subjects. This encompasses several different issues, the main ones being who the authority is legitimate over and who ought to be delegating to what authority. In the previous section, we explained who has to obey the authority, and we touched on a paradigmatic case, as well as some cases that do not fall into that standard paradigm. We now need to say something about what authority the subjects should obey.

The question of what authority the subjects should obey is not a trivial one. This is because there might be many political authorities that satisfy the two conditions of legitimacy. This might be optimistic, but several political authorities might present their subjects with adequate plans that help them respond to reasons and give their subjects a basic justification for their plans. However, this does not mean that we have a duty to delegate to all authorities that satisfy the two conditions of legitimacy we presented.⁷² The main point here is that we have good reasons to think that one should always be looking to delegate to some authority, and which authority one ought to delegate to depends on one's political status and where one is at a given moment in time. Generally, one always ought to obey and delegate to the political authority that has jurisdiction over the territory one is a citizen of, resident of or refugee in. A resident of the UK, for instance, will always have to delegate her planning power to the UK government, provided the government in authority in the UK at that time is putting forward adequate, publicly justified plans. Even if this UK resident goes on holiday in another country, or moves to another country, she still has to delegate to the UK government (usually in the form of paying taxes if she is a UK citizen). A related point to make is that, if we find ourselves transiting through or temporarily residing in a different state than the one we are citizens of, we should also delegate our planning powers to the authority that has power over this new

⁷² However, we might be under a duty not to interfere with the delegation activity of other legitimate authorities.

territory we are in. This means that there might be cases where people need to be responsive to and delegate to multiple political authorities at the same time. As an Italian citizen residing in England, for instance, I need to be responsive both to the Italian government and the UK government. Given that I spend 95% of my time in England each year, I have to pay more attention to the plans the UK government imposes on me. However, I cannot forget my duties to the Italian government, and I need to ensure I fulfil my obligations to my home country. Moreover, if I decided to move to the Netherlands for two months, I would still classify as an Italian citizen and a UK resident, as my visa highlights. For this reason, I would have to delegate to three different political authorities: the UK government, the Italian government and – for the two months I am there – the Netherlands government.

To summarise these points, we should only obey and delegate to political authorities that satisfy the two conditions of legitimacy we laid out. However, this does not mean we need to obey *any* authority that satisfies them. We should be responsive and delegate our planning powers to the political authority that has jurisdiction over the territory we live and reside in. Some complicated cases of dual citizenship or different place of citizenship and residence, for instance, mean that some people ought to delegate to different authorities at the same time.

This is how political authorities are legitimate according to the delegation theory.

7.4 Objections to normative delegation

In Chapter 6, I established the conditions under which political authorities are legitimate for the delegation theory. In this chapter, I explained that normative delegation is a better ground for political legitimacy than standard, voluntary delegation. In this section, I address two objections that are usually advanced against accounts that ground something like legitimacy on a duty to consent or delegate on the part of the citizens. The first one deals with delegation

being a spare wheel in accounts that argue for its normative rather than voluntary form. The second objection claims that it is wrong for individuals or states to treat others as if they consented or delegated tasks to them when that consent or delegation did not occur.

The first objection argues that accounts of normative delegation, just as accounts of normative consent,⁷³ make delegation a spare wheel that does not pull any normative weight. I said before that, in my account of normative delegation, the subjects are treated as if they delegated their social planning powers to the state even though actual delegation on their part did not occur. When the subjects have reasons to delegate (i.e., the state puts forward the kind of adequate plans we talked about, and it publicly justifies its operations), they are treated as if they did. The subjects are, in other words, put in a default position that amounts to them still being obligated to obey the law and treated as such regardless of whether they delegated or not. But does this not make delegation a spare wheel? With normative delegation, it seems as though the normative heavy lifting in justifying the authority's actions and grounding its legitimacy is done by the conclusive reasons to obey the state that the citizens have. The authority produces plans whose adequacy is confirmed by a public justification on the part of the authority. The problems that these plans need to solve are very important and the citizens have a duty to coordinate their behaviour with that of fellow citizens to solve them, which generates a duty to delegate to an authority that is capable of solving those problems should those problems be too difficult for individual practical reasoning. The normative weight does not seem to rest on the shoulder of delegation in this case. At this point, it is natural to ask why we need delegation at all. If it doesn't have any normative weight, we could just do without it.

In response to this challenge, let us consider what delegation does and stand for in theories of normative delegation. In such theories, delegation, or better, citizens who delegate

⁷³ Accounts of normative consent have been heavily criticised. For some objections see Frank (2016); Koltonski (2013); and Manson (2013).

to the state, just as consent in accounts of normative consent (Estlund, 2008), are a paradigm, an ideal that societies strive or should strive to achieve. These theories take the presence of some conditions to warrant treating the citizens as if they had delegated to the authority. In an ideal setting, virtually everyone would delegate under those specific conditions, and the state would then be authorised to act in certain ways as a result. In a non-ideal setting, not everyone delegates despite those favourable conditions being present. The state, however, still has important functions to perform and, to get them done and help its subjects, it puts them in a default position of having delegated (when only some of them have delegated). This means that the state treats everyone as having obligations to obey the law. In this case, the state is forced, in a way, to put everyone under this default position of having delegated to achieve the important societal goals it is supposed to achieve. By treating everyone as being obligated to obey the law, the state tries to prevent possible disobedience that would negatively impact those who have actually delegated their planning powers to it. This, though, does not detract value from delegation itself. On these theories, the state keeps delegation in mind in everything that it does. It considers cases when the citizens would delegate before it treats them as having done so: it does not put them under the default position of delegation when the conditions that would make it legitimate are not present. In this case, delegation is a normative ideal that states strive to achieve and refer to in their operations. Delegation's normative force is still present, and it becomes evident in instances where people actively dissent or express their disapproval towards actions carried out by the state.

Dissent or disapproval signal to the state that that person or group is disagreeing with the policy or directive issued by the state, or with the actions the state has performed. By dissenting, citizens signal to the state that they should not be taken as having delegated their social planning powers anymore. Dissent, then, pulls the state's subjects out of that default position of being treated as having delegated that the state puts them into when it has important

functions to perform. In this respect, delegation still plays an important normative role in normative accounts, albeit in reverse. Voluntary delegation grounds the state's legitimacy and, as a result of it, states act in certain ways to reach societal goals. When it comes to normative delegation, states still require delegation to discharge important functions and, in its absence, they assume that everyone has delegated and treat them as such. Only when people actively express their dissent, delegation pulls them out of that default position and makes it so that the state's actions are not performed on their behalf. So, in instances of voluntary delegation, individuals delegate first and then the state acts accordingly as a result of that act of delegation. In cases of normative delegation, the state assumes at first that everyone has delegated (the default position) and treats them as such unless they express their dissent to the state's policies or actions. Delegation has normative force in both cases of voluntary and normative delegation. That normative force is simply activated at different times depending on the account we endorse.

So, in normative accounts, consent and delegation still maintain the normative force they exhibit in voluntary accounts. Furthermore, accounts of legitimacy based on normative delegation provide us with a more realistic theory of legitimacy than accounts of voluntary delegation or consent. Classic (political) consent theories have been criticised because they require the consent of all or most group members for an authority or its policies to be legitimate. Normative delegation accounts do not fall prey to the same criticism, as citizens are all treated as if they delegated as a default when favourable conditions allow (i.e., urgent tasks, adequate plans and public reasoning). This enables the state to discharge important functions without the hindrance of lack of consent or delegation on the part of all or most of its subjects. It is difficult to see how most citizens in political societies would consent or delegate to the authority. It is not that far-fetched, however, to think that the state can treat them as if they delegated (until or unless they dissent).

Something similar is at play in opt-out accounts of post-mortem organ procurement (Saunders, 2011), where governments assume that all their citizens are organ donors unless they ask to be taken out of the donor registry. Governments could wait for people to register as donors (which is what goes on in opt-in systems of post-mortem organ procurement) or put everyone in the default position of being a donor. Choosing the latter would maximise the number of people on the organ donor registry, thus allowing the state to save more people who need an organ transplant. Countries that adopt opt-out systems of organ procurement do not neglect consent to donating organs and people's consent still plays a fundamental normative role in these situations. The state treats everyone as organ donors by default. However, the moment someone express their dissent to donating organs by asking to be taken out of the donor registry, the state is not allowed to harvest that person's organs after their death. Consent's normative force manifests itself when people dissent. The state can make a similar assumption about its subjects having delegated their planning powers. It can operate under the assumption that everyone delegated and treat everyone as having obligations to obey the law. However, the moment people express their dissent, delegation's normative force manifests itself by pulling them out of the default state of having delegated. Granted, it is much more difficult to establish what happens when citizens make clear they did not delegate their planning powers to the state. They cannot simply be taken out of the list of citizens who abide by the law. It is my intention to explore dissent in a different project. However, it is plausible to say that the state can still treat its subjects as having delegated their planning powers and that this does not deprive delegation of its normative force. Dissent can be used to let the state know that delegation did not occur, and this is when its normative power 'activates'.

It is realistic for the state to make assumptions about its subjects and treat them in certain ways as a result of those assumptions. Delegation, as we explained, entails the transfer of our agential powers to someone else so that that person or institution can act on our behalf. But is

it wrong to violate someone else's agential powers by assuming they have delegated and treating them as if they did? If we consider the case of consenting to things such as sexual intercourse or granting access to our property, it is quite evident how wrong and damaging it can be when we assume that we have someone's consent for those things without actually having it. As we said before, consent is in place to protect us as individuals and our bodily integrity. When someone refuses to consent to someone else accessing their body or possessions, they erect a barrier that works to protect them from unwanted contact or interactions with others. If I assumed that my friend consented to me borrowing her laptop, I would do something wrong by grabbing it, as the absence of her authorisation to use it means that I am actively stealing it. Even worse, when one assumes that someone has consented to sexual intercourse with them and acts accordingly, they would do something deeply wrong in acting as if the other party has consented. Our intuition in these cases is that bodily integrity, alongside the sphere of property and possessions that surrounds us, seems to be a barrier that we can hardly ever transgress. Consent, at least in some cases, seems to be more individualistic in nature than political delegation which, as we argued, entails the transferring of agential powers that are social and collective at their core. This makes a difference in what we are allowed to do with someone else's agential powers.

To see if in instances of normative delegation, and even normative consent, it is less wrong to take over someone's agential powers, let us consider delegation in small-scale and large-scale cases. In small-scale settings, it is problematic if someone exercises our agential powers without our authorisation. This is because we could have likely done something ourselves, but someone else has hijacked our agential powers by acting on our behalf without us asking for or agreeing to it. In my previous example of an accountant doing my taxes, in the absence of my authorisation, it would be wrong if a random person started going through my files and putting together a tax application for me. This is not just because that individual is

violating my privacy by accessing personal, sensitive information about me. It is also wrong because that person is taking away the opportunity for me to freely exercise my agential powers. It might be convenient that she does the taxes for me, and it might even save me time and avoid some unnecessary stress. However, I did not choose that this person alleviates my worries and workload, and something I could have done myself was carried out by someone else.⁷⁴

Things are different in large-scale settings, where the social aspect of our agential powers becomes more evident. In political societies, for instance, there needs to be centralised planning for everyone to be able to exercise their agential powers, and it is difficult to act without it. In small-scale cases, there is always an alternative available to delegating our agential powers to someone else: we can just do it ourselves. If delegation falls through, we have the option of exercising our own agential powers and completing a task. In large-scale cases, this alternative to delegation is unavailable because of the inadequacy of individual agential powers in solving complex social coordination problems. In these cases, we only have two options: either the state performs these kinds of tasks for us, or they do not get done, since it is improbable that individuals would step up and complete them on their own. If I do not delegate to the state in these cases, the state will not perform important functions that affect the entire community. If those important functions are not discharged, individuals cannot exercise their agential powers as the community itself does not function without that social coordination. So, when I refuse to delegate to the state, assuming that the state is there to perform important functions through adequate plans, I leave other people and myself stranded. I also violate other agents' agential powers because refusing to delegate does not allow the state to coordinate us in a way that allows everyone to exercise those powers. When it comes to large-scale contexts and those

⁷⁴ It might even be detrimental to our development at a young age if people continuously hijack our capacity to make decisions and do things on our behalf just because they can.

instances of delegation (or consent) that involve (many) others, the negative implications that not delegating has on others are obvious.

When the state has important, urgent tasks to perform, then, it is not wrong of it to act as if its subjects delegated their agential powers to it. By assuming that they delegated, the state can carry out its functions and avoid the inaction that would cause more agential powers to be violated. If only a handful of people refuse to delegate, this is not going to make a big, long-lasting impact on the state's actions, and it is also likely to go unnoticed. If a sufficient number of people refuse to delegate, though, that kind of free riding can impair the state's activities and prevent others (who presumably have delegated) from fully exercising their agential powers. And we need to find ways to prevent that. As Jonathan Wolff put it:

So it is likely that a situation in which we can free ride is a situation in which many other people can do the same, and large-scale free riding would undermine any benefits that the state has to offer. Soon there would be nothing to free ride on (Wolff, 1991, p. 159).

People can and sometimes even have a duty to dissent. If free riding (by not delegating to the state) for the sake of free riding is to be avoided, expressing one's dissent for inadequate or grossly ineffective social plans is important to ensure our autonomy is protected. We have a duty to coordinate with others in political societies but if our institutions do not put forward adequate plans, that duty does not give rise to the related duty to delegate our agential powers to the state.⁷⁵

⁷⁵ I intend to explore the subjects' duty to dissent, the treatment of dissenting portions of the population and how all that impacts the state's legitimacy in a post-doctoral research project.

Chapter 8 – Conclusion

My primary aim in this thesis was to defend the possibility of legitimate political authority. I understood legitimate political authority, *de jure* authority, as having the right to rule over those who are in its jurisdiction. This right to rule is accompanied by a duty to obey the authority's directives and laws on the part of the subjects. I also understood political authority as a kind of practical authority, that is, authority that influences people's conduct by providing them with reasons for action.

I argued that political authorities and their subjects can be engaged in hierarchical shared activities as members of large-scale groups (nation-states). These complex instances of acting together under the authority's directions are sustained by plans created by the authority to coordinate its subjects' actions. What individuals and authorities have in common is the ability to plan their present and future conduct. For this reason, I investigated the possibility of political legitimacy as a product of large-scale, hierarchical agency, where the agential power to create plans is transferred from the subjects to the authority. I then defended what I call the *Delegation Theory of the Legitimacy of Political Authority*, according to which political authorities are legitimate when they provide their subjects with adequate plans that help them respond to reasons, and when they give a basic public justification for their plans. When these two conditions are met, agents ought to delegate their planning powers to the authority.

However, those who are interested in investigating the possibility of legitimate authority encounter a challenge: Robert Paul Wolff's autonomy-authority problem. According to Wolff (1970), the moral autonomy of individuals is incompatible with how political authorities operate. Obeying a political authority means doing what the authority tells you to do *because* the authority tells you to do it. However, autonomy or self-legislation entails having a duty to

give yourself moral commands. Wolff thinks that it is not possible to do something *because* the authority tells you to do it and to self-legislate at the same time. So, authority and autonomy cannot coexist for him. Wolff's concern that legitimate political authority is incompatible with individuals being and acting autonomously deeply affects those accounts of authority that assign a duty to obey authoritative directives to individuals. The duty to obey authoritative directives without assessing their merit is thought to be incompatible with autonomous decision-making. According to the delegation theory of political legitimacy, the subjects have a duty to obey the authority. So, the theory needed to address Wolff's concerns to explain the relationship between legitimate authorities and their subjects. It became clear that the legitimacy question, a normative question regarding the conditions political authorities need to meet to be legitimate, could not be seriously answered without addressing the autonomy question, a normative question about whether individual autonomy and political authority can coexist.

Before I could answer the legitimacy question, I needed to demonstrate that legitimate political authority and autonomy are not always incompatible. Legitimate authority is valuable because it possesses the resources, time and knowledge to solve problems that arise in nation-states, and it can do so by adhering to certain planning standards. *De facto* authority can discharge the same functions. However, there are ways of planning and solving societal issues that are better than others. We might think that an authority that consistently adheres to certain standards of planning is justified in its claim of having authority over others. Suppose Wolff is right in saying that legitimate authority impairs autonomy and that we are happy to say that authority can be legitimate under some circumstances. There is something quite not right about exercising legitimate authority over non-autonomous individuals. Even if we do not subscribe to the view that we have a duty to be autonomous as Wolff does, there is still value in

autonomous decision-making that would be lost if legitimate authority came about at the expense of our autonomy.

So, my secondary aim in this thesis was to defend the claim that legitimate political authority and autonomy can be compatible. I divided the thesis into two parts. Part 1, which includes Chapters 2, 3, 4 and 5, is dedicated to addressing Wolff's autonomy-authority problem. Part 2, which comprises Chapters 6, 7 and 8, is dedicated to the legitimacy question and provides an account of what conditions political authorities need to meet to be legitimate.

After outlining the debate and giving an overview of the thesis in the Introduction (Chapter 1), I examined Wolff's autonomy-authority problem in Chapter 2. I explained why it is important to respond to his anarchist challenge and surveyed two possible ways to solve the problem he presents. Looking at the definitions of autonomy and authority Wolff relies on, I contended that the Strong View on authority, according to which obeying an authority means doing what the authority tells you to do *because* the authority tells you to do it, is the right understanding of how political authorities operate. I also modified Wolff's strong autarkic view on autonomy and showed that the tension between autonomy and authority remains even if we adopt a weaker conception of autonomy. I then analysed how three theories of the legitimacy of political authorities respond to Wolff's challenge. I concluded that Joseph Raz's account of legitimate authority gets close to giving us a full, satisfactory picture of how autonomy and authority can coexist. So, I closely followed Raz's account in solving the autonomy-authority problem. I provided an account of legitimate authority with some modifications with respect to reconsidering authoritative plans, which the subjects must treat as exclusionary.

In Chapters 3, 4 and 5 I defended a three-stage model of practical reasoning to reconcile autonomy and legitimate authority. A detailed analysis of Wolff's argument revealed a *prima facie* general tension between autonomy and shared agency. The supposed clash between

autonomous decision-making processes and following authority stems from treating directives as exclusionary. However, the theory of action revealed that any plan, not just plans made by authorities, is exclusionary in nature. Plans must be treated as exclusionary after settling the matter on what to do to effectively guide our conduct. This raised the question of whether we lose our autonomy when we follow plans we made in the past, or when we act with others, with or without authority relations in the group. Building on influential work in the theory of action, the three-stage model analysed all relevant moments of decision-making, from acting individually to acting with others and complying with authority. It showed our autonomy is compatible with treating plans as exclusionary in these contexts.

In Chapter 3, I argued that treating plans as exclusionary facilitates our decision-making processes by reducing the range of options available to us to choose from. By streamlining our decision-making, plans help us respond to our reasons. Moreover, by acting as metaphorical repositories of reasons, plans made in the past reconnect us with the reasons that went into the creation of the plans when we act on them in the future. So, they not only help us respond to our reasons but also enhance our reason-responsiveness capacity by allowing us to act on our past reasons far in the future. This enables a complex organisation of thoughts and actions towards future goals.

In Chapters 4 and 5, I showed that plans perform all the functions above even when we are engaged in egalitarian shared activities and hierarchical shared activities. In contexts of egalitarian shared agency, when we act with others as equals, shared plans do not impair our autonomy. The group's reasons can become our reasons if we voluntarily join the group or come to identify with it. Moreover, shared plans can instrumentally help us respond to our reasons when coordinating and organising our conduct with others is necessary for us to do what we have reasons to do. Similarly, in contexts of hierarchical shared agency, authority's plans can also help us respond to reasons by providing the necessary social order and

coordination needed for us to reach our goals. Furthermore, when the authority transmits reasons to us via its plans, abiding by those plans enhances our reason-responsiveness capacity.

The three-stage model of practical reasoning presented in Chapters 3, 4 and 5 offers a view on autonomy that is not based on total self-reliance. This view is compatible with acting together and delegating tasks to others, including authorities.

In Part 2 of the thesis (Chapters 6, 7 and 8), I addressed the legitimacy question. Once I established that autonomy and legitimate authority could be compatible, I needed to investigate the conditions of legitimate political authority. In Chapter 6, I contended that political authorities are legitimate when they provide their subjects with adequate plans that help them respond to reasons, and when they give a basic public justification for their plans. The desideratum for adequate plans is that they are sensitive to solving societal problems and, throughout the chapter, I specified the complexity of this desideratum. It is important that political authorities publicly justify their plans. This gives their subjects a chance to evaluate the adequacy of authoritative plans if needed and decide whether they should reopen deliberation about what to do if necessary. I believe the requirement that the authority publicly justifies its plans gives more freedom to the subjects, as they can decide for themselves when it is appropriate to reconsider following authoritative plans. Public reasoning, then, makes treating the authority's plans and reasons as exclusionary less rigid than it would otherwise be.

When the two conditions above are met, the subjects ought to delegate their social planning powers to the authority. In Chapter 7, I introduced and started analysing normative delegation, which encompasses the subjects' duty to delegate planning to the authority. We have a duty to delegate because societal problems are pressing to solve, and the state is in a better position than us to solve them. The authorities we delegate our social planning powers to are legitimate only when they produce adequate plans that they publicly justify. Only when

these two conditions are met the subjects' duty to delegate to the authority will be activated. When that happens, agents are treated as if they have delegated, regardless of whether they have done so.

In Chapter 7, I also explained that the duty to delegate our social planning powers to the authority is related to a more general duty to coordinate our actions with other people's actions. Finally, at the end of Chapter 7, I defended normative delegation from two objections that might be raised against it.

This analysis of delegation in general and normative delegation in particular only began to scratch the surface of a complex and multi-faceted avenue of research in political philosophy and action theory. A more comprehensive defence of the Delegation Theory would show the implications of the theory on the workings of representative democracy. Representative democratic governments are often seen as delegates that defer to their subject's judgement to act and make decisions (Pitkin, 1967). Representative democracy, then, seems to be a natural practical application of my theory of political legitimacy. A future research project should analyse representative democracy through the lenses of the theory of action, as well as applying and adapting the Delegation Theory to representative democratic government.

Another interesting aspect of the theory I did not have the space to adequately analyse is the potential incompatibility of delegation and public reasoning. Public reasoning is often invoked as a solution to the representatives abusing the powers their citizens delegated to them. Public reasoning forces representatives to justify their behaviour to their subjects who, in turn, can keep them in check (Christiano, 2005; Richardson, 2002). However, delegation entails the outsourcing of tasks to others and a division of labour between two parties. Public reasoning on the part of the representatives threatens to compromise both. We outsource tasks to others to save time and energy, or because we do not have the skills needed to perform them. If we

need to check what the other party is doing, why bother delegating at all? Would it not be more efficient and less time-consuming to just do it oneself? Furthermore, when delegating a task, we place its execution into someone else's hands. This division of labour is violated by public reasoning. Delegates need to be able to do their job without interference. Demanding that they continuously explain themselves violates the boundaries of that division of labour. So, outsourcing tasks to others and division of labour are embedded in delegation. At the same time, delegates should not have unlimited powers and delegation needs to be somehow regulated. Public reasoning is necessary to regulate delegates but seems incompatible with delegation itself. A more comprehensive analysis of delegation will need to address these concerns.

In addition to that, it would also be fruitful to explore the connection between delegation and responsibility, and the implications of that connection on debates in the ethics of protest. De Smet *et al.*'s account of responsibility (2016) highlights the connection between delegation and responsibility in matters of climate change. They argue that delegating tasks to someone else does not relieve us from responsibility related to what others do on our behalf. We still need to see to it that those tasks are completed and that the authority is not abusing its powers. If the subjects retain responsibility for the actions authorities do on their behalf, it means that when the authority produces immoral plans, the subjects are (perhaps partially) responsible for those wrongdoings. One way in which they can sever that connection between delegating powers to the authority and being responsible for the authority's immoral action is by protesting and actively dissenting. This leads me to believe that the subjects might have a duty to protest that is activated when the authority violates the terms of representation or delegation.

The Delegation Theory, then, has the potential to make contributions to debates on representative democracy, public reasoning and the ethics of protest. I hope to employ the

interdisciplinary methodology used throughout the thesis to illuminate these matters in the future.

Bibliography

- Adams, N. P. 2018. The Relational Conception of Practical Authority. In *Law and Philosophy*, 37(5): 549-575.
- Aghion, P., & Tirole, J. Formal and Real Authority in Organisations. In *Journal of Political Economy*, 105(1): 1-29.
- Alexander, L. 2014. The Ontology of Consent. In *Analytic Philosophy*, 55(1): 102-113.
- Alvarez, M. 2017. Reasons for Action: Justification, Motivation, Explanation. In *The Stanford Encyclopedia of Philosophy* (Winter 2017 Edition), Edward N. Zalta (ed.), URL=<https://plato.stanford.edu/archives/win2017/entries/reasons-just-vs-expl/>
- Anderson, E. 2010. *The Imperative of Integration*. Princeton, NJ: Princeton University Press.
- Angst, L. & Borowiecki, K. J. 1993. Delegation and Motivation. In *Theory and Decision*, 76(3): 363-393.
- Anscombe, G. E. M. 1963. *Intention*. 2nd ed. Ithaca: Cornell University Press.
- Anscombe, G. E. M. 1979. On the Source of the Authority of the State. In *Ratio: An International Journal of Analytic Philosophy*, 20: 1-28.
- Archard, D. 2010. The Obligations and Responsibilities of Parenthood. In Archard D. & Benatar, D. (eds.), *Procreation and Parenthood: The Ethics of Bearing and Rearing Children*. Oxford: Oxford University Press, 103-127.
- Arneson, R. 2003. Defending the Purely Instrumental Account of Democratic Legitimacy. In *Journal of Political Philosophy*, 11(1): 122-132.

- Arneson, R. 1993. Democratic Rights at National and Workplace Levels. In Copp, D; Hampton, J & Roemer, J. E. (eds.). *The Idea of Democracy*. Cambridge: Cambridge University Press.
- Asgeirsson, H. 2013. Vagueness and Power-Delegation in Law: A Reply to Sorensen. In Freeman, M., and Smith, F. (eds.), *Current Legal Issues: Law and Language*. Oxford: Oxford University Press.
- Bates, S. 1972. Authority and Autonomy. In *The Journal of Philosophy*, 69(7): 175-179.
- Bazargan-Forward, S. 2022. *Authority, Cooperation and Accountability*. New York, NY: Oxford University Press.
- Beauchamp, T. L. 2009. Autonomy and Consent. The Nature of Consent. Miller F. & Wertheimer, A. (eds.). *The Ethics of Consent: Theory and Practice*. Oxford: Oxford University Press.
- Beran, H. 1987. *The consent theory of political obligation*. London: Croom Helm.
- Bird, C. 2014. Coercion and Public Justification. In *Politics, Philosophy, and Economics*, 13: 142–214.
- Blackman, L. 2010. Bodily Integrity. In *Body and Society*, 16(3): 1-9.
- Boettcher, J. 2015. Against the Asymmetric Convergence Model of Public Justification. In *Ethical Theory and Moral Practice*, 18 (1): 191–208.
- Bogg, A. L. 2002. In Defence of Correlativity: Authority, Autonomy and Mistaken Legal Directives. In *Ratio Juris*, 15(1): 84-96).
- Bordieu, P. & Robinson, K. 1985. Delegation and Political Fetishism. In *Thesis Eleven*, 10(1): 56-70.

- Bratman, M. E. 1999a. *Faces of Intention. Selected Essays on Intention and Agency*. Cambridge: Cambridge University Press.
- Bratman, M. 1987. *Intentions, Plans and Practical Reason*. Cambridge, MA: Harvard University Press.
- Bratman, M. E. 1999b. *Intention, Plans and Practical Reason*. Stanford, CA: CSLI Publications.
- Bratman, M. E. 2009. Intention, Practical Rationality, and Self-Governance. In *Ethics*, 119(3): 411-443.
- Bratman, M. E. 2018. *Planning, Time, and Self-Governance: Essays in Practical Rationality*. Oxford: Oxford University Press.
- Bratman, M. E. 2014. *Shared Agency: A Planning Theory of Acting Together*. Oxford: Oxford University Press.
- Bratman, M. E. 2022. *Shared and Institutional Agency. Toward a Planning Theory of Human Practical Organisation*. Oxford: Oxford University Press.
- Bratman, M. E. 1992. Shared Cooperative Activity. In *The Philosophical Review*, 101(2): 327-341.
- Bratman, M. E. 1983. Taking Plans Seriously. In *Social Theory and Practice*. 9(2-3): 271-287.
- Broome, J. 2001. Are Intentions Reasons? And How Should We Cope with Incommensurable Values. In Morris, C. W. & Ripstein, A. (eds.). *Practical Rationality and Preference: Essays for David Gauthier*. Cambridge: Cambridge University Press, pp. 98-120.
- Brower, B. W. 1994. The Limits of Public Reason. In *The Journal of Philosophy*, 91(1): 5-26.
- Brunero, J. 2007. Are Intentions Reasons? In *Pacific Philosophical Quarterly*, 88(4): 424-444.

- Butterfill, S. A., & Apperly, I. A. 2013. How to Construct a Minimal Theory of Mind. In *Mind & Language*, 28(5): 606-637.
- Carr, C. R. 1983. The Problem of Political Authority. In *The Monist*, 66(4): 472-486.
- Christiano, T. 2005. Democracy and Bureaucracy. In *Philosophy and Phenomenological Research*, 71(1): 211-211.
- Cibik, M. forth. Tacit Consent and Political Legitimacy. In *European Journal of Political Theory*.
- Cohen, J. 2009. Truth and Public Reason. In *Philosophy & Public Affairs*, 37(1): 2-42.
- Collins, S. 2019. *Group Duties: Their Existence and Their Implications for Individuals*. Oxford: Oxford University Press.
- Dagger, R. & Lefkowitz, D. 2014. Political Obligation. In *The Stanford Encyclopedia of Philosophy* (Fall 2014 Edition), Zalta, E. N. URL = <https://plato.stanford.edu/archives/fall2014/entries/political-obligation/>.
- Davidson, D. 1963. Actions, Reasons and Causes. In *The Journal of Philosophy*, 60(23): 685-700.
- Davidson, D. 1980. *Essays on Actions and Events*. Oxford: Oxford University Press.
- Denters, B., & Klimovsky, D. 2017. Participation and Democratic Accountability: Making a Difference for Citizens. In Schwab, C.; Bouckaert, G. et. al. *The Future of Local Government in Europe: Lessons from Research and Practice in 31 Countries*. Baden-Baden: Nomos Verlagsgesellschaft mbH, pp. 79-100.

- De Smet, A.; Peeters, W., and Sterckx, S. 2016. The Delegated Authority Model Misused as a Strategy of Disengagement in the Case of Climate Change. In *Ethics & Global Politics*, 9(1): 292-299.
- Dodds, S. & Jones, K. 1989. Surrogacy and Autonomy. In *Bioethics*, 3(1): 1-17.
- Donahue, S. 2019. Public Justification and the Veil of Testimony. In *The Journal of Political Philosophy*, 28(4): 378-396).
- Dougherty, T. 2015. Yes Means Yes: Consent as Communication. In *Philosophy & Public Affairs*, 43(3): 224-253.
- Douglas, T. 2016. Duties to Rescue: Individual, Professional and Institutional. In *Journal of Medical Ethics*, 42(4): 207-208.
- Downey, L. 2021. Delegation in Democracy: A Temporal Analysis. In *The Journal of Political Philosophy*, 29(3): 305-329.
- Dworkin, R. 1986. *Law's Empire*. Cambridge, MA: The Belknap Press of Harvard University Press.
- Ehrenberg, K. 2011. Joseph's Raz Theory of Authority. In *Philosophy Compass*. 6(12): 884-894.
- Edmundson, W. 1998. *Three Anarchical Fallacies*. Cambridge: Cambridge University Press.
- Enoch, D. 2013. The Disorder of Public Reason. In *Ethics*, 125: 141–176.
- Estlund, D. 2008. *Democratic Authority: A Philosophical Framework*. Cambridge: Cambridge University Press.
- Feinberg, J. 1989a. Autonomy (pp. 27-51). In Feinberg, J. *The Moral Limits of the Criminal Law, Volume 3: Harm to Self*. New York, NY: Oxford University Press.

- Feinberg, J. 1989b. Consent and its Counterfeits (pp. 172-188). In Feinberg, J. *The Moral Limits of the Criminal Law, Volume 3: Harm to Self*. New York, NY: Oxford University Press.
- Frank, N. 2016. Against Normative Consent. In *Journal of Social Philosophy*, 47(4): 470-487.
- Furner, J. 2010. Tacit Consent Without Political Obligation. In *Theoria: A Journal of Social and Political Theory*, 57(124): 54-85.
- Gardiner, S. M. 2011. *A Perfect Moral Storm: The Ethical Tragedy of Climate Change*. New York, NJ: Oxford University Press.
- Gauthier, D. P. 1986. *Morals by Agreement*. New York: Oxford University Press.
- Gilbert, M. 2006. *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*. Oxford: Oxford University Press.
- Gilbert, M. 2014. *Joint Commitment: How We Make the Social World*. New York: Oxford University Press.
- Gilbert, M. 2015. Joint Commitment: What It Is and Why It Matters. In *Phenomenology and Mind*, 9:18-26.
- Gilbert, M. 1996. *Living Together: Rationality, Sociality and Obligation*. Lanham, MD: Rowman and Littlefield.
- Gilbert, M, 2000. *Sociality and Responsibility: New Essays in Plural Subject Theory*. Lanham, MD: Rowman and Littlefield.
- Gilbert, M. 1999. Walking Together: A Paradigmatic Social Phenomenon. In *Midwest Studies in Philosophy*. 15(1): 1-14.

- Green, L. 2004. Associative Obligations and the State. In Burley, J. (ed.), *Dworkin and His Critics*. Oxford: Blackwell.
- Green, L. 1989. *The Authority of the State*. Oxford: Clarendon Press.
- Harman, G. 1986. *Change in View*. Cambridge, MA: MIT Press.
- Hart, H. L. A. 1982. *Essays on Bentham: Studies in Jurisprudence and Political Theory*. Oxford: Oxford University Press.
- Hart, H. L. A. 1961. *The Concept of Law*. Oxford: Clarendon Press.
- Hess, K. M & Orthmann, C. 2008. *Introduction to Law Enforcement and Criminal Justice*. CENGAGE Delmar Learning.
- Hobbes, T. 2003 [1651]. *Leviathan: a critical edition*. Rogers, G. A. J and Schuhmann, K. eds. Bristol: Thoemmes.
- Horton, J. 2006. In Defence of Associative Political Obligations: Part I. In *Political Studies*, 54(3): 427-442.
- Hurd, H. 1991. Challenging Authority. In *The Yale Law Journal*. 100: 1611-1677.
- Hurd, H. 1996. The Moral Magic of Consent. In *Legal Theory*. 2(2): 121-146.
- Hume, D. 1965 [1748]. Of the Original Contract. In MacIntyre, A. (ed.). *Hume's Ethical Writings*. 1965. London: University of Notre Dame Press.
- Kant, I. 1996, [1785]. The Groundwork of the Metaphysics of Morals. In Guyer, P. & Wood, A. (eds.), *The Cambridge Edition of the Works of Immanuel Kant*. Cambridge: Cambridge University Press.
- Keller, S. 2013. *Partiality*. Princeton: Princeton University Press.
- Kleinig, J. 1982. The Ethics of Consent. In *Canadian Journal of Philosophy*, 8: 91-118.

- Kleinig, J. 2009. The Nature of Consent (pp. 3-22). In Miller, F. & Wertheimer, A. (eds.). *The ethics of consent: theory and practice*. New York, NY: Oxford University Press.
- Kolodny, N. 2014. Rule Over None I: What Justifies Democracy? In *Philosophy and Public Affairs*, 42(3): 195-229.
- Koltonski, D. 2013. Normative Consent and Authority. In *Journal of Moral Philosophy*, 10(3): 255-275.
- Korsgaard, C. M. 1989. Personal Identity and the Unity of Agency: A Kantian Response to Parfit. In *Philosophy & Public Affairs*, 18(2): 101-132.
- Kutz, C. 2000. Acting Together. In *Philosophy and Phenomenological Research*, 61(1): 1-31.
- Lackey, J. 2020. *The Epistemology of Groups*. New York, NY: Oxford University Press.
- Lee, I. 2015. Filial Duty as the Moral Foundation of Caring for the Elderly: Its Possibility and Limitations. In Fan, R. (ed.), *Family-Oriented Informed Consent: East Asian and American Perspectives*. Springer Cham, pp. 125-136.
- Lewis, D. 1983. Survival and Identity. In Lewis, D. *Philosophical Papers Volume I*. Oxford: Oxford University Press.
- List, C. & Pettit, P. 2006. Group Agency and Supervenience. In *The Southern Journal of Philosophy*, 44(S1): 85-105).
- List, C. & Pettit, P. 2011. *Group Agency. The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford University Press.
- Locke, J. 2016 [1689]. *Second treatise of government and A letter concerning toleration*. Goldie, M. ed. Oxford: Oxford University Press.

- Luban, D. 2009. Authority and Responsibility in International Criminal Law. In Tasioulas, J., and Besson, S. (eds.), *The Philosophy of International Law*. Oxford: Oxford University Press.
- Ludwig, K. 2017. *From Plural to Institutional Agency: Collective Action II*. Oxford: Oxford University Press.
- Ludwig, K. 2020a. The Social Construction of Legal Norms. In Garcia-Godinez, M., Mellin R., and Tuomela, R. *Social Ontology, Normativity and Law*. De Gruyter: 179-208.
- Ludwig, K. 2020b. What Is Minimally Cooperative Behaviour? In Fiebich, A. (ed.). *Minimal Cooperation and Shared Agency*. Cham: Springer, pp. 9-40.
- Lupia, A. 2001. Delegation of Power: Agency Theory. In Smelser, N. J., and Baltes, P. B., *International Encyclopedia of the Social and Behavioral Sciences*. Amsterdam: Elsevier.
- Manson, N. C. 2013. Normative Consent is Not Consent. In *Cambridge Quarterly of Healthcare Ethics*, 22(1): 33-44.
- Mill, J. S. 1859 [2001]. *On Liberty*. Kitchener, ON: Batoche Books Limited. Accessed online on 05/09/2023 at <https://socialsciences.mcmaster.ca/econ/ugcm/3ll3/mill/liberty.pdf>
- Nozick, R. 1974. *Anarchy, State and Utopia*. Oxford: Blackwell.
- Parfit, D. 1971. Personal Identity. In *Philosophical Review*. 80(1): 3-27.
- Perry, S. R. 1989. Second-Order Reasons, Uncertainty and Legal Theory. In *Southern California Law Review*, 62, 945-963.

- Peter, F. 2007. Rawls' Idea of Public Reason and Democratic Legitimacy. In *Politics and Ethics Review*, 3(1): 129-143.
- Pitkin, H. F. 1967. *The Concept of Representation*. Berkeley, CA: Berkeley University Press.
- Prusak, B. G. 2013. *Parental Obligations and Bioethics: The Duties of a Creator*. New York, NY: Routledge.
- Puryear, S. 2021. Consent By Residence: A Defence. In *European Journal of Political Theory*, 20(3): 529-546.
- Qvortup, M. 2013. *Direct Democracy: A Comparative Study of The Theory and Practice of Government by The People*. Manchester: Manchester University Press.
- Railton, P. 1986. Moral Realism. In *Philosophical Review*, 95: 163-207.
- Raz, J. 1990. (Ed.). *Authority*. New York: New York University Press.
- Raz, J. 2009. *Between Authority and Interpretation: On the Theory of Law and Practical Reason*. Oxford: Oxford University Press.
- Raz, J. 1995. *Ethics in the Public Domain: Essays in the Morality of Law and Politics*. Oxford: Clarendon Press.
- Raz, J. 1979. *The Authority of Law: Essays on Law and Morality*. Oxford: Oxford University Press.
- Raz, J. 1986. *The Morality of Freedom*. New York, NY: Oxford University Press.
- Raz, J. 2006. The Problem of Authority: Revisiting the Service Conception. In *Minnesota Law Review*, 90: 1003-1044.
- Rawls, J. 1971. *A Theory of Justice*. Cambridge, MA: Belknap Press.
- Rawls, J. 1957. Justice as Fairness. In *Journal of Philosophy*, 52(22): 653–662.

- Rawls, J. 1958. Justice as Fairness. In *Philosophical Review*, 67 (2): 164–194.
- Rawls, J. 2001. *Justice as Fairness: A Restatement*. Cambridge: Harvard-Belknap Press.
- Richardson, H. S. 2002. *Democratic Autonomy: Public Reasoning About the Ends of Policy*.
Oxford: Oxford University Press.
- Ripstein, A. 2000. Three Duties to Rescue: Moral, Civil and Criminal. In *Law and
Philosophy*, 19(6): 751-779.
- Ritchie, K. 2020. Minimal Cooperation and Group Roles. In Fiebich, A. (ed.). *Minimal
Cooperation and Shared Agency*. Cham: Springer, pp. 93-109.
- Roth, A. S. 2020. Collective Responsibility and Entitlement to Collective Reasons for Action.
In S. Bazargan-Forward & D. Tollefsen (eds.), *Routledge Handbook for Collective
Responsibility*. NY: Taylor & Francis.
- Roth, A. S. 2017. Entitlement to Reasons for Action. In Shoemaker, D. (ed.), *Oxford Studies in
Agency and Responsibility, Vol. 4*. Oxford: Oxford University Press, pp. 75-92.
- Roth, A. S. 2014. Prediction, Authority and Entitlement in Shared Activity. In *Noûs*, 48(4):
626-652.
- Roth, A. 2004. Shared Agency and Contralateral Commitments. In *The Philosophical Review*,
113(3): 359-410.
- Rousseau, J. 1712-1778 [2019]. *The Social Contract and Other Later Political Writings*. Edited
and Translated by Victor Gourevitch. 2nd ed. Cambridge: Cambridge University Press.
- Saunders, S. 2009. Normative Consent and Opt-out Organ Donation. In *Journal of Medical
Ethics*, 36(2): 84-87.
- Scanlon, T. M. 1998. *What We Owe To Each Other*. Cambridge, MA: Belknap Press.

- Scheffler, S. 2018. Membership and Political Obligation. In *Journal of Political Philosophy*, 26(1): 3-23.
- Schroeder, T. 2020. Desire. *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition), Zalta, E. N. (ed.), URL=
<https://plato.stanford.edu/archives/sum2020/entries/desire/>.
- Schroeder, M. 2007. *Slaves of the Passions*. Oxford: Oxford University Press.
- Schwitzgebel, E. 2021. Belief. *The Stanford Encyclopedia of Philosophy* (Winter 2021 Edition), Zalta, E. N. (ed.), URL =
<https://plato.stanford.edu/archives/win2021/entries/belief/>.
- Searle, J. 1990. Collective Intentions and Actions. In Cohen, P. R. & Pollack, M (eds.), *Intentions in Communication*. Cambridge, MA: MIT Press, pp. 401-415.
- Searle, J. 1983. *Intentionality. An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press.
- Searle, J. 1995. *The Construction of Social Reality*. Free Press.
- Shapiro, S. J. 2004. Authority. In Coleman, J.; Himma, K. E. & Shapiro, S. J. (eds.). *The Oxford Handbook of Jurisprudence and Philosophy of Law*.
- Shapiro, S. J. 2011. *Legality*. Cambridge, MA: Harvard University Press.
- Shapiro, S. J. 2014. Massively Shared Agency. *Rational and social agency: the philosophy of Michael Bratman* (pp. 257-289). In Vargas, M. and Yaffe, G. eds. New York, NY: Oxford University Press.
- Shildrick, M. 2010. Some Reflections on the Socio-cultural and Bioscientific Limits of Bodily Integrity. In *Body and Society*, 16(3): 11-22.

- Simmons, A. J. 1996. Associative Political Obligations. In *Ethics*, 106(2): 247-273.
- Simmons, A. J. 1998. “Denisons” and “Aliens”: Locke’s Problem of Political Consent. In *Social Theory and Practice*, 24(2): 161-182.
- Simmons, A. J. 2001. *Justification and Legitimacy. Essays on Rights and Obligations*. New York, NY: Cambridge University Press.
- Simmons, A. J. 1993. *On The Edge of Anarchy: Locke, Consent, and the Limits of Society*. Princeton, NJ: Princeton University Press.
- Simmons, A. J. 2009. Political Obligation and Consent (pp. 305-326). In Miller, F. & Wertheimer, A. (eds.). *The ethics of consent: theory and practice*. New York, NY: Oxford University Press.
- Simmons, A. J. 1976. Tacit Consent and Political Obligation. In *Philosophy and Public Affairs*, 5(3): 274-291.
- Singer, P. 1972. Famine, Affluence and Morality. In *Philosophy and Public Affairs*, 1(3): 229-243.
- Smith, M. 1994. *The Moral Problem*. Oxford: Basil Blackwell.
- Smith, P. 1990. The Duty to Rescue and The Slippery Slope Problem. In *Social Theory and Practice*, 16(1): 19-41.
- Sorensen, R. 2001. Vagueness has No Function in Law. In *Legal Theory*, 4: 385-415.
- Sorley, W. R. 1901 [2015]. *Recent Tendencies in Ethics: Three Lectures to Clergy Given at Cambridge*. Victoria, AL: Leopold Classic Library.
- Star, D. & Delmas, C. 2011. Three Conceptions of Practical Authority. In *Jurisprudence*, 2(1): 143-160.

- Stark, C. A. 2000. Hypothetical Consent and Justification. In *Journal of Philosophy*, 97(6): 313-334.
- Steinberger, P. 2004. *The idea of the state*. Cambridge: Cambridge University Press.
- Stuifbergen, M. C. & Van Delden, J. J. M. 2011. Filial Obligations to Elderly Parents: A Duty to Care? In *Medicine, Health Care & Philosophy*, 14(1): 63-71.
- Swank, O. H., & Visser, B. 2006. Motivating Through Delegating Tasks or Giving Attention. In *Journal of Law, Economics and Organisation*, 23(3): 731-742.
- The Brussels Times. 2020. Belgium breaks own record for longest period without government [Online]. [accessed on 26/08/2023]. Available from: <https://www.brusselstimes.com/124777/belgium-breaks-own-record-for-longest-period-without-government>
- Tollefsen, D. 2015. *Groups as Agents*. Malden, MA: Polity.
- Tuomela, R. & Miller, K. 1988. We-intentions. In *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 53(3): 367-389).
- Tuomela, R. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.
- Viehoff, D. 2014. Democratic Equality and Political Authority. In *Philosophy and Public Affairs*, 42(4): 337-375.
- Viens, A. M. 2014. *The Right to Bodily Integrity*. Lund Humphries Publishers.
- Wall, S. 2006. Democracy, Authority, and Publicity. In *Journal of Political Philosophy*, 14(1): 85-100.

- Wallerstein, S. 2015. Delegation of Powers and Authority in International Criminal Law. In *Criminal Law and Philosophy*, 9(1): 123-140.
- Way, J. 2015. Reasons as Premises in Good Reasoning. In *Pacific Philosophical Quarterly*, 98, 251-270.
- Weber, M. 1919 [1970]. Politics as a Vocation (pp. 77-128). In Gerth, H. H. & Wright Mills, C. (eds.). *From Max Weber: Essays in sociology*. London: Routledge.
- Wellman, C. H. 1996. Liberalism, Political Legitimacy, and Samaritanism. In *Philosophy and Public Affairs*, 25: 211-237.
- Westen, P. 2004. *The Logic of Consent: The Diversity and Deceptiveness of Consent as a Defense to Criminal Conduct*. London: Routledge.
- Westlund, A. C. 2018. Relational Autonomy and Practical Authority. In Geravaso, P. (ed.) *The Bloomsbury Companion To Analytic Feminism*. Bloomsbury Academic, pp. 375-393.
- Williams, G., & Chin, G. 2000. The Failure of Citizens's Initiated Referenda Proposals in Australia: New Directions for Popular Participation? In *Australia Journal of Political Science*, 35(1): 27-48.
- Wolff, J. 1991. What Is the Problem of Political Obligation? In *Proceedings of the Aristotelian Society*, 91: 153-169.
- Wolff, R. P. 1970. *In Defense of Anarchism*. Berkeley & Los Angeles: University of California Press.
- Yaffe, G. 1995. Velleman on Intentions as Reasons for Action. In *Analysis*, 55(2): 107-115.