

Washington University in St. Louis

Washington University Open Scholarship

All Theses and Dissertations (ETDs)

Spring 4-27-2014

Evidential Modern Political Authority

Nathan Adams

Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

Recommended Citation

Adams, Nathan, "Evidential Modern Political Authority" (2014). *All Theses and Dissertations (ETDs)*. 1214.
<https://openscholarship.wustl.edu/etd/1214>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS

Department of Philosophy

Dissertation Examination Committee:

Christopher Heath Wellman, Chair

Eric Brown

Julia Driver

Charlie Kurth

Frank Lovett

Evidential Modern Political Authority

by

Nathan Paul Adams

A dissertation presented to the
Graduate School of Arts and Sciences
of Washington University in
partial fulfillment of the
requirements for the degree
of Doctor of Philosophy

May 2014

St. Louis, Missouri

© 2014, N. P. Adams

Table of Contents

Acknowledgements	iii
Introduction	1
1. Authority and Anarchism	14
2. Evidential Modern Political Authority	61
3. How Authorities Bind	99
4. From Expertise to Authority	141
5. Towards Evidential Governmentalism	180
Works Cited	226
Appendix	229

Acknowledgements

My most profound and sincere thanks to Kit Wellman; without his support, trust, and encouragement this would not exist. His incisive and constructive critiques challenged me at every turn, ultimately making both my thinking and writing about these complex issues immeasurably better. Thanks also to all the members of the philosophy department, especially Isaac Wiegman, whose strength of mind and character sharpened me. Finally, the funding and institutional support of Washington University in St. Louis made this possible.

To
JC, LJ, CE
for
always pushing
always supporting
always encouraging
always loving
and
knowing what matters

Introduction

The task of this dissertation is to answer one of the fundamental challenges in modern political philosophy: the question of under what conditions a modern government can attain genuine moral authority over its citizens. It is clear that governments claim moral authority because they claim that they can issue laws that citizens may be punished for disobeying. Punishment, however, involves harsh treatment of the sort that persons generally have standing moral rights against, like being imprisoned. These rights stand in the way of permissibly punishing people for acting in ways you disagree with, yet that is exactly what the government claims to be able to do. Governments claim that laws somehow change the moral standing of citizens such that they no longer have a right against being punished for acting in ways that entail breaking the law.

On the other hand, it seems quite clear that many actual modern governments lack the moral standing to issue morally binding commands. It strains credulity to claim that a brutal dictator like Kim Jong-Un has the moral permission to punish North Koreans for dissenting political speech, for example. There are bad laws and bad regimes. A good account of political authority provides one important and intuitive way of distinguishing governments that actually morally bind their citizens from governments that do not.

Although understanding the conditions under which a government has political authority are clear, it is common in philosophy today to endorse philosophical anarchism, the position that no actual government has or can plausibly attain genuine moral authority. While anarchists generally admit that there are often moral reasons to do as the law says,

especially under reasonably just regimes, they deny that any government has the power to issue laws that bind citizens' and ipso facto change their moral standing with respect to particular acts.

While philosophical anarchism is in some respects counterintuitive, it is also theoretically robust. When political thought moved away from the divine right of kings and began focusing on the rights of individual citizens and how those rights constrain rulers, political authority became a much more difficult issue. Since then the problem of political authority has most often been approached through the lens of consent. Consent is the most obvious way that an impermissible act can be made permissible, and it is clear that if a government could secure actual consent from its citizens, then punishment wouldn't wrong them. However, anarchists have persuasively argued that no actual government has secured such consent and that the extant attempts to establish governmental authority by appealing to consent in its non-actual forms, for example tacit consent or hypothetical consent, are flawed.

The goal of the dissertation, then, is to explore the idea of political authority and to propose and defend a novel account of political authority. I argue that this new account shows that some modern governments, particularly liberal democracies, have genuine moral authority over most of their citizens. This authority contributes to an explanation of how those governments govern permissibly. However, my account differs from traditional accounts of political authority, and those differences matter for our understanding of the relationship between citizens and the law. In the end we can respect many of the intuitions motivating philosophical anarchism, and many of its arguments, without embracing the counterintuitive position that no government has the power to issue morally binding laws.

In the end a plausible and informative picture of the relationship between citizens and the law emerges. All citizens have a natural duty of justice to establish, maintain, and support the institutions that constitute the basic structure of society in a way that meets the demands of justice. In order to set up and maintain these institutions over time while meeting the demands of justice, the government must coordinate action on a mass scale by establishing a system of punishment that doesn't wrong its citizens. The government has expertise relative to almost all their citizens about how to establish and support these institutions, meaning the citizen is much more likely to contribute to just institutions by following the law than if she tried to establish and support institutions based on her own judgment. Due to this, when the government issues commands, citizens' evidence about how to best discharge their natural duty of justice, and so how not to wrong their co-citizens, decisively changes. This change makes them vulnerable to punishment for acting otherwise, and so is sufficient to ground a system of permissible punishment. When the government has sufficiently widespread expertise about setting up and maintaining the basic institutions of society in a way that treats all citizens fairly, laws morally bind citizens and their actions may be coercively coordinated in order to meet our shared natural duty of justice.

The five chapters that follow this introduction make the argument for this picture. It begins with a detailed characterization of authority and claims that authority must be understood more broadly than is generally assumed in the debate over political authority. This broader notion allows us to see that what is required for the sort of authority governments claim is evidential authority, rather than traditional authority. A close look at the nature of evidence leads to a better understanding of the way in which commands bind

subjects, and so in what sense citizens are morally bound to do as the law dictates. The main condition on evidential authority is epistemic expertise, so we enter the social epistemology literature and explore the idea that governments could plausibly achieve the relevant expertise due to their institutional nature. Over the course of this discussion I add four additional conditions on genuine evidential authority, each responding to concerns over when it is reasonable for citizens to defer to authorities. Finally, I argue that some modern governments plausibly meet these conditions and so have genuine evidential authority over most of their citizens in most domains, denying philosophical anarchism. The rest of the introduction more fully details this argument.

1. Authority and Anarchism

The main task of the first chapter is to characterize the concept of authority. The starting point is the notion of command: authorities issue commands, speech acts that are intended to bind subjects by giving them a preemptive and content-independent reason. Preemptive reasons are contrasted with merely additive reasons, and the idea is that preemptive reasons replace (or preempt) some of the subjects' other reasons rather than merely being added to the overall balance of subjects' reasons. Content-independent reasons are reasons that have force due to their provenance rather than to their content; when a mother commands her child to go to her room, the reason for the child to obey is that the mother has the appropriate standing, not that the content of the command was in some sense correct. Had a stranger issued the same directive to the child, the child wouldn't be bound because the provenance of the command was wrong even though the content of the command was the same. This characterization of authority traces through Thomas Hobbes

and Jeremy Bentham to the twentieth century and the philosophies of H. L. A. Hart and Joseph Raz. These theories have in common a rationalist approach to authority, characterizing authority by the sort of reasons commands constitute for subjects.

The most important argument of chapter one focuses on the moral standing the authority has to issue commands. The traditional notion of authority is of authority as the right to command. Here 'right' refers to a Hohfeldian advantage. When one agent has a right against another, the other has a correlative duty. Thus if authority is the right to command, subjects have a correlative duty to obey. This is in many ways a very intuitive and plausible concept of authority. If citizens have a duty to obey the law, the idea that they can be punished for breaking the law makes intuitive sense. However, the traditional notion of authority is too narrow.

The idea of authority as the right to command cannot capture some paradigmatic exercises of authority. We have to expand our understanding of authority past the idea of a right to command and focus on a different Hohfeldian advantage, namely on a power. When one agent has a power over another, then she can alter the other's normative standing at her discretion. This modified understanding of authority shows that agents can have authority that effects change in the standing of subjects without changing what they have a duty to do. Thus we see the possibility of authority that isn't traditional authority, and the possibility that non-traditional authority could be attained even though traditional authority cannot, as the anarchists have shown.

Finally, I outline Raz's service conception of authority. The service conception details sufficient conditions for genuine authority and is the basis of my new account. The main condition on genuine authority for Raz is the normal justification thesis, which says

that one agent has authority over another when the subject would likely better conform to her reasons if she tried to follow the directives of the authority than if she tried to follow her own judgment. Thus authority is in service to the subject, only binding her because she will better conform to her own preexisting reasons.

2. Evidential Modern Political Authority

With the concept of authority clarified and a conception of authority on offer, I then turn to narrowing the focus of the dissertation to modern political authority in particular. Modern political authority is characterized in six ways that distinguish it from authority in general and that show why modern political authority is especially difficult to attain. In sum, the problem of modern political authority is this: how could one agent have the power to command millions or even billions of incredibly diverse individuals, over the course of their entire lives, about the most important domains of their lives, in such a way that punishment for disobedience is morally permissible, yet without their consent. If attaining genuine moral authority over even one other person in any condition is difficult, as it seems to be, then attaining modern political authority seems supremely difficult.

The focus of the rest of the chapter is the idea of punishable authority, or authority that issues commands subjects can be permissibly punished for disobeying. This is the defining and most morally worrying feature of modern political authority. The question is what kind of authority is punishable authority. At first glance, traditional authority looks exactly right because it creates duties, so subjects act wrongly if they disobey. This isn't right though, as I show with a series of thought experiments. To fully understand moral vulnerability to punishment, we have to distinguish between wrongdoing and culpability.

Someone is vulnerable to punishment only if they are culpable, yet agents can be culpable without acting wrongly and can be nonculpable despite acting wrongly. Culpability is grounded in subjective wrongdoing rather than objective wrongdoing: an agent is culpable if she acts in a way that would be objectively wrong if her beliefs or evidence were true, rather than if she acts in a way that is objectively wrong.

Thus governments require the power to change citizens' subjective duties if they are to permissibly punish citizens for breaking the law. While traditional accounts can capture this change without much difficulty, our broadened understanding of authority as a power instead of a right shows that governments could have authority that changes subjective duties without changing objective duties. In this case they have evidential rather than traditional authority. In some situations, when subjects' evidence changes so does what they are prima facie objectively vulnerable to punishment for doing, regardless of whether their objective duties have changed. This is important because it opens up the possibility of a kind of political authority that anarchists haven't addressed and that modern governments can perhaps attain.

There is a worry with this thought from the outset, however. Changing evidence merely changes subjective duties, but authority must change subjects' objective standing. Further, it must change objective standing in such a way that subjects' actions can be coordinated in order to achieve political aims. It appears that changes in epistemic reasons cannot accomplish this task. However, a more nuanced understanding of evidence and its relationship to objective vulnerability to punishment answers this worry. I can coordinate action and make a decisive difference in subjects' objective moral standing by making

individuals' objectively vulnerable to punishment. Evidential authority really is practical authority.

3. How Authorities Bind

Traditional authority runs into a dilemma when considering nonconsensual authorities that issue mistaken commands. Authorities are agents and all agents are fallible, so all actual authorities will sometimes issue mistaken commands. However, it cannot be the case that mistaken commands necessarily fail to bind. If this were the case, authority would be an empty and useless idea, for subjects would simply have reason to do as their preexisting reasons entail and whether an authority commanded the subject to act in some way is irrelevant. Further, the fact that commands give content-independent reasons entails that some mistaken commands bind, for whether a command is mistaken is an issue of content. The problem for traditional authority is that mistaken commands are supposed to entail a moral duty to do as commanded, but the preexisting reasons also entail a duty to do not as commanded (given that the command is mistaken). This contradicts the plausible principle of "ought implies can".

The most obvious response to this dilemma is to claim that the authority's command somehow removes the countervailing reasons. Thus Raz claims that countervailing reasons are excluded, which is a special kind of second-order defeat. Excluded reasons, regardless of their weight, can no longer function as reasons for the subject. However, getting clear on the notion of exclusion is difficult. Raz endorses a motivational interpretation of exclusion, but this interpretation should be rejected because it requires an implausible understanding of our motives and because it is dissonant with out

commonsense understanding of what it means to obey an authority. Instead we should understand exclusion under the justificatory interpretation. The justificatory interpretation shows that consensual authority excludes by canceling the force of certain reasons, while nonconsensual, evidential authority excludes by undercutting the force of certain reasons. Undercutting defeat is a function of reliability, which coheres well with our understanding of evidence, Raz's examples of exclusion, and the normal justification thesis.

This leads us to the first condition on authority, which is an analog to the normal justification thesis. I call it the expertise condition. Experts are individuals who are likely to be less mistaken and likely to be mistaken less than the layperson because they have extensive training in a particular domain. It is especially important and relevant that they understand the second-order justificatory structure of the domain, so they can modify their beliefs in light of new evidence, can apply their beliefs to new situations, and so forth. This means that the reliability of their judgments trumps the reliability of non-experts' judgments: when an expert issues a judgment, the rational response is to believe what the expert says precisely because she is an expert.

4. From Expertise to Authority

Chapter four details three further conditions on genuine evidential modern political authority. First is the acceptance condition, which says that an authority must issue directives with practical intent, with the intention that the subject take the command to give sufficient reason to act on its basis in the practical context under consideration. This distinguishes authority from the mere issuance of expert judgments and bridges the gap

from expert directives that give reason to believe to practical commands that give reason to act.

However, this move makes it appear that in order to have moral authority you must also have moral expertise. This is because an expert about a non-moral domain does not necessarily know when her expertise is relevant to the subject's reasons in a particular context and so when she should claim the subject has a reason to accept her judgment. But if moral expertise is required for moral authority, it looks very implausible that modern governments could have any form of moral authority. The answer to this concern is the precedence condition. Even non-experts sometimes, even often, know when non-moral expertise is relevant in a given moral context. While moral expertise is (perhaps) sufficient for this knowledge, it isn't necessary. What's required for moral authority grounded in non-moral expertise is knowledge that the non-moral expertise is precedent in a particular practical context, i.e. that acting on that domain is also acting all-things-considered.

It might also appear that moral authority requires moral virtue. Authority is not about judgment but about issuing directives, and one way to fail to issue binding directives is to issue directives when one shouldn't. This is something like the deficit we see in the old story of the boy who cried wolf. The problem here isn't whether the purported authority is expert, but whether she exercises her authority by issuing directives in the appropriate manner. This returns to the idea that command is a communicative speech act that requires a particular relationship between two agents. This relationship must be one of trust, for otherwise deference to another is rationally mystifying. Trust is best understood as relying on the goodwill and competence of another by granting them discretionary powers over important interests. The more important the interests in

question and the more powers you grant them, the more vulnerable you make yourself to them and the more trust is required.

Finally, I address whether it is even *prima facie* plausible that governments could attain the levels of expertise and trustworthiness required for political authority. Governments have extensive powers over some of the most important interests in citizens' lives, so require an extensive amount of expertise and trustworthiness that they do not at first glance seem apt to attain. The answer to this problem lies in the institutional nature of modern governments. Institutions are able to aggregate the expertise of many individuals in order to come to expert judgments. The expertise we are concerned with is an emergent property of institutions rather than of particular individuals in the government. The same is true of trustworthiness: institutions are able to incorporate explicit, public rules that ground goodwill in a way that individuals cannot. This is not to say that institutions necessarily do this, just that well-constructed institutions can raise the expertise and trustworthiness of the government to levels much higher than we would expect when considering the expertise and trustworthiness of the individuals that make up modern governments.

5. Towards Evidential Governmentalism

In the final chapter I apply the theory's five conditions to the case of Western democracies and argue that they plausibly have genuine evidential authority over most of their citizens in most domains. This is sufficient authority to achieve governments' ends and denies philosophical anarchism. First, expertise. I argue that some modern governments are expert about the demands of justice, understood in a Rawlsian vein, over most of their

citizens. This is not the claim that modern governments are absolute or universal experts about justice, but merely that they are expert relative to most of the population. This is especially true because justice requires contextual expertise, as the demands of justice vary from context to context. While some citizens may be experts about justice in the abstract, they likely lack the contextual expertise that is required to make all but the most basic demands of justice determinate in a particular context.

Further, once we see that the natural duty of justice requires mere equal vulnerability to punishment, not universal obedience to the law, we see another way traditional theories have gone wrong. Simmons rightly objects to Rawls that having a duty to obey the law is not necessary for justice to be achieved; an individual act of law-breaking has no significant effect. But universal vulnerability to punishment is a different matter, for without that the law is arbitrary and can't support and maintain just societal institutions.

Governments clearly claim practical authority, so the acceptance condition is met. It is also intuitively clear that the natural duty of justice is important enough to meet the deference condition. And the governments' expertise, both in justice and in non-moral domains, is clearly relevant to individual citizens' duty to create, uphold, and maintain just societal institutions, meeting the precedence condition.

The remaining condition is the trustworthiness condition. I address this condition, and bolster the case for the relative, contextual expertise of modern governments, by looking at democracy. Democratic institutions are important firstly because they guarantee (in some ways, although certainly not all) that citizens' interests are accounted for in the judgments of the government. Poor enough decisions will result in politicians not

being reelected and so changes in how the government functions. This contributes greatly to the goodwill component of trustworthiness.

Democratic institutions also contribute to the competence and ultimately the expertise of modern governments. This is because feedback mechanisms like elections constrain the judgments of governments in ways that enhance competence. Further, there is recent social science evidence that under the right conditions, cognitive diversity is more important for the competency of group decisions than individuals' levels of competence. Finally, the fact that many political questions are morally indeterminate means that we can make decisions on procedural rather than substantive grounds. These reflections show why my view that the government requires expertise to have genuine moral authority does not entail that experts should rule.

Some modern governments are trustworthy experts about meeting the demands of justice with respect to most of their citizens in most domains, giving them genuine evidential modern political authority. While this entails that philosophical anarchism is false, the account presented here respects the anarchist's motivations while allowing for a nuanced, plausible, and intuitive understanding of the relationship between citizens and the governments that largely structure our lives in modern societies. In the end we better understand the conditions for justified governance, and so as citizens have stronger and clearer grounds for holding governments accountable to our interests and to the achievement of justice.

1 – Authority and Anarchism

One of the central questions of political philosophy is whether governments have moral authority over their citizens. If the government has moral authority, then its commands constitute changes in citizens moral standing: simply because the government has made some act illegal changes citizens' moral relations to that act. Philosophical anarchists deny that any actual government has genuine (or *de jure* or justified) moral authority, in part because the idea that governments as we know them can simply create moral duties *ex nihilo* seems preposterous.¹ In contrast to philosophical anarchists, governmentalsists claim at least some actual governments have genuine moral authority, in part because of the intuitively plausible idea that government and the issuance of morally binding laws is a necessary part of achieving great moral goods, such as justice.

The debate between philosophical anarchists and governmentalsists has a long history, with entrenched arguments on both sides. This debate, however, has been myopic. It focuses entirely on one kind of moral authority, what I call *traditional* authority. My sense of the literature and theorists' general attitudes is that philosophical anarchists have won the debate over traditional authority, so anarchism has become the default position on the question of political authority. I basically agree with this trend. The anarchists have established, to my mind, that governments don't have (and realistically can't attain) traditional authority. Traditional governmentalsism, the view that governments have

¹ See, e.g., A. John Simmons, "The Anarchist Position: A Reply to Klosko and Senor", *Philosophy and Public Affairs* 16 (1987): pp. 269-70. I take this to be the standard understanding of philosophical anarchism, but others define it in a broad variety of ways, including views about political legitimacy, sovereignty, and other governmental attributes. I will use it simply as the position that denies any actual government has genuine moral authority.

traditional authority, is false.

Philosophical anarchism can no longer be the default position once we get past the myopic focus on traditional authority. Once we see that there are types of moral authority other than traditional authority, we see that governments could have moral authority despite lacking traditional authority. In this dissertation I argue for a version of governmentalism that attempts to capture this middle ground. I call it *evidential governmentalism*. In this chapter I more fully articulate the anarchist and governmentalist positions, clarify the concept of authority, and begin to defend the myopia charge.

1. Anarchism and Governmentalism

Philosophical anarchism, as I have said, is the position that no government has genuine moral authority. It is contrasted with political anarchism, which goes beyond the claim that governments lack moral authority to some practical conclusions about how citizens should relate to the government. The political anarchist thinks the government is unnecessary, should be resisted, and perhaps even overthrown. Philosophical anarchists are often not political anarchists because they believe there are good reasons to do as the government says and keep the government in place despite the fact that the government lacks moral authority.² Political anarchism looks especially implausible if philosophical anarchism is false, so if I can establish a version of governmentalism and deny philosophical anarchism, political anarchism will fall as well. As such my focus will be exclusively on philosophical anarchism, and I will use ‘anarchism’ to mean philosophical anarchism.

² Although there are some reasons to doubt that philosophical anarchism without political anarchism isn't possible. For example, see Christopher Heath Wellman and A. John Simmons, *Is There a Duty to Obey the Law?* (Cambridge: Cambridge UP, 2005): pp. 26-29.

Philosophical anarchism can be divided into conceptual philosophical anarchism and empirical philosophical anarchism.³ Conceptual anarchists argue that governments lack moral authority because the very idea of moral authority is conceptually incoherent. Conceptual anarchists are a rarity, and to my mind have been plausibly and convincingly answered.⁴ The more common position is empirical anarchism, according to which governments lack moral authority because no actual government meets the conditions for such authority. According to the empirical anarchist, there are possible worlds where a government could meet the conditions on genuine moral authority (like an imaginary world where governments were able to secure actual universal consent). The problem is that, as the world is, no government does meet those conditions.

Within empirical philosophical anarchism we can make further distinctions. According to the empirical anarchist, the fact that no government has genuine authority is a contingent fact. If empirical circumstances changed sufficiently, an empirical anarchist would be a governmentalist. The question, then, is how close actual governments are to meeting the conditions on genuine authority. At one limit we might think actual governments are extremely close to meeting those conditions, so a few tweaks would change us from anarchists to governmentals. At the other limit we might think actual governments are extremely far away from meeting those conditions, so far away that nothing that looks like a government as we conceive of them could plausibly meet the conditions. Our anarchism would then be much less contingent, to the point that we might think no government could have moral authority in any plausibly-realizable state of the

³ See Simmons, "The Anarchist Position", pp. 269-70 for this contrast.

⁴ The only two I know of are Robert Paul Wolff, *In Defense of Anarchism* (New York: Harper & Row, 1970) and Heidi M. Hurd, "Challenging Authority", *The Yale Law Journal* 100 (1991): pp. 1611-77. See Simmons, "The Anarchist Position", n. 2 for a refutation of Wolff. I address Hurd's similar argument in Appendix A.

world.

This more robust, less contingent empirical anarchism describes the position of most anarchists, including empirical anarchism's exemplar, A. John Simmons. Simmons rejects conceptual anarchism because he thinks that universal actual consent could ground a government's genuine moral authority. Universal actual consent, though, is the only way to genuine political authority on Simmons' account. While we can imagine a possible world where something like a modern government was able to secure universal actual consent, no actual government has done this. Further, it is plausible that no actual government could do this. The logistics are a nightmare, there is the ever-present possibility of contrarians who would not consent simply because everyone else has, and many people in the world today are not in a position to give morally binding consent. So for Simmons, the possible world where a government has genuine authority is very far away from the actual world and any plausibly-realizable government must lack moral authority. Simmons' more robust empirical anarchism can be put thusly: all actual governments lack genuine moral authority and in no foreseeable, plausibly-realizable near future would any government be able to attain it. This is the most common, well-defended, and plausible version of anarchism. In addition, it is the most practically relevant version of anarchism, as it focuses on the sort of government that could plausibly affect the wellbeing of people in the world today. It will be my main focus for these reasons.

In contrast to anarchism as I've defined it is governmentalism. Most minimally, governmentalism is the claim that at least one government has moral authority over at least one citizen in at least one domain. The term 'governmentalism' is my own, but this position has various names in the literature, including 'statism'. I think government is best

thought of as the group of individuals filling roles in a political institution while the state is a particular kind of political institution.⁵ Drawing the distinction in this way shows that it is preferable to think of the authority of governments rather than the authority of states, and so counts in favor of using ‘governmentalism’ rather than ‘statism’. First, authority rests with agents, not with institutions. We can make sense of obeying an individual who fills an institutional role because they fill that institutional role, but the idea of obeying an institution as such is at least more mysterious.

Second, states are only one kind of political institution, especially if we take ‘state’ to refer to something like modern nation-states. There are other forms of political organization that establish non-state political institutions, for example global institutions. If we frame the discussion of political authority in terms of state authority, we risk ignoring the possibility that non-state governments could have genuine political authority. As I am a political cosmopolitan and believe that justice in our modern world demands some non-state political institutions, I especially want to highlight the possibility of genuine non-state political authority.

Third, using the term ‘statism’ threatens to burden our discussion of political authority with expectations that are particular to state authority. A good example of this is the claim modern nation-states paradigmatically, and perhaps conceptually, make to *perfectly general* authority. That is, they claim authority over all individuals within their territory and with respect to all domains over those individuals. If we use the term ‘statism’, then, it looks reasonable to demand that a good account of political authority be

⁵ We could also think of the government as the collective agent that emerges from individuals acting in institutional roles. I am sympathetic to this construction but don’t wish to get into the complicated and controversial issues surrounding collective agency.

able to justify authority that is perfectly general. But almost everyone in this debate doesn't think political authority as such must be perfectly general with respect to individuals or domains. We shouldn't expect an account of political authority to give criteria for the precise sort of authority that actual modern nation-states claim. Instead, as I explain below, there are certain features of modern states that a good account of political authority must capture but there are others that can be discarded, like perfect generality. So instead of 'statism', I'll use 'governmentalism' to refer to the position that governments have genuine authority.

Governmentalists must be precise about what sort of moral authority they think actual governments have or can plausibly attain. The minimalist governmentalist position claims that there is only one actual government that has genuine moral authority over only one individual in only one domain. While this position is a conceptual possibility, as far as I know nobody takes it. On the other extreme, the maximalist governmentalist position claims that all actual governments have perfectly general authority of the sort they claim. As with the minimalist position, this is unusual, especially in philosophy. Its basis would presumably have to be something like the claim that effective power is sufficient for genuine moral authority. As a view about moral authority this would be bizarre, licensing the idea that even the world's worst dictators have moral authority over the people they oppress.

The most common governmentalist position lies between the maximalist and minimalist versions. On this view, the governments of some Western democracies, paradigmatically the Scandinavian countries but including others, have moral authority of a sort that is very similar to the authority they actually claim. Many nations in the world,

probably due to injustice on their part, lack any moral authority over their subjects.

The evidential governmentalism I'll argue for in this dissertation is a version of this common, middle position. My view will also endorse the claim that many governments entirely lack moral authority over their subjects. But those governments that do have genuine authority will have it to a lesser extent on my account than on standard governmentalist approaches. Following Joseph Raz, I'll argue that authority is first and foremost a three-part relation between the authority, a subject, and a domain.⁶ A has authority over S in D. It is only as a summary of individual authority relations that we can speak of perfectly general authority, authority over all individuals (in a territory, paradigmatically) with respect to all domains. An authority's *jurisdiction* is defined by the domains and individuals she has authority over.

On this radically individualized, piecemeal approach to authority, a government could have moral authority over S1 only with respect to domains D1, D2, D5, and D7 while simultaneously having moral authority over another S2 only with respect to domains D1, D3, and D6. This is nothing like perfectly general authority and my approach will not completely endorse the kind of sweeping authority the standard governmentalist often claims to capture. Scandinavian governments will have authority over more citizens and more domains than other governments, even others that have some genuine moral authority over some citizens and some domains, like the American government. That said, the more restricted the government's authority, the less plausible governmentalism looks. In order to achieve its ends, the government must coordinate action between most of its citizens and so must have authority over most of its citizens. So the interesting and

⁶ Joseph Raz, *The Morality of Freedom* (Oxford: Oxford UP, 1986). [MF]

controversial governmentalist claim is not simply that some modern governments have some kind of moral authority, but that they have sufficient moral authority to actually achieve their aims.

While my approach will ground evaluations of actual governments similar to other approaches, I will argue genuine political authority is of a different nature than governmentalist have generally assumed, and so the conditions for attaining genuine political authority quite different. Even under the new conditions, though, we will see that some modern governments can have sufficiently broad authority to achieve their aims. This new picture of authority also changes, and I hope clarifies, our understanding of the appropriate relationship between citizens and the law.

2. Commands as Speech Acts

The disagreement between the anarchist and the governmentalist hinges on the idea of authority. In order to evaluate whether governments have authority, and over which individuals and which domains, we need an account of authority. This section begins to present such an account, although I will continue to build and modify the account throughout the course of the dissertation. The traditional understanding of authority claims that authority is the *right to command*.⁷ Eventually I argue that “the right to command” is too narrow an understanding of authority, but this characterization is an intuitive and sufficient starting point. Let’s begin with the notion of command.

⁷ For example, see Robert Paul Wolff, “The Conflict Between Authority and Autonomy”, in ed. Joseph Raz, *Authority* (New York: NYU Press, 1990): p. 20; G.E.M. Anscombe, “On the Source of the Authority of the State”, in ed. Raz, p. 144; Raz, MF, p.23, and see n.1 on that page for further instances of this view.

First and foremost, commands are speech acts.⁸ A paradigmatic example is a mother exercising her parental authority and commanding her child, “Go to your room!”. What makes this speech act a command? In its most everyday usage, a command is simply any speech act that takes the form of an imperative. But in the context of authority, ‘command’ is generally a technical term. The best way to understand the difference between the everyday and technical senses of “command” is through the lens of Austin’s theory of speech acts.⁹

Austin distinguishes between three different aspects of a speech act: the locutionary aspect, the illocutionary aspect, and the perlocutionary aspect. Imagine a child coming indoors on a cold day and her mother asking the question “Is the door shut?”. The locutionary aspect of this utterance is interrogative. It can be answered in the affirmative or negative, depending on whether the door is shut. But often when parents ask questions of their children, they mean to do more than simply ask a question. This extra element is the illocutionary aspect of the utterance. When the mother asks whether the door is shut, she’s really also requesting that the child shut the door if it’s not shut. The request is illocutionary; it’s not reflected in the structure of the utterance but is still intended to convey the mother’s request that the door be shut. The utterance “Please shut the door” is a request in both its locutionary and illocutionary aspects. Finally, the perlocutionary aspect of an utterance depends on how the listener responds to the utterance. If the illocutionary request of “Is the door shut?” is successfully conveyed to the child, she shuts the door. Shutting the door was the perlocutionary response to the illocutionary request of

⁸ See H. L. A. Hart, “Commands and Authoritative Reasons”, in *Essays on Bentham* (Oxford: Clarendon Press, 1982): pp. 243-268. Hart traces the focus on commands as speech acts to Bentham, ultimately arguing that Bentham’s philosophy of language was unable to account for the nature of command.

⁹ J. L. Austin, *How to Do Things with Words* (Oxford: Oxford UP, 1955).

the interrogative locution of “Is the door shut?”.

The everyday usage of ‘command’ focuses on the locutionary aspect of utterances. Here ‘command’ simply refers to any speech act with the locutionary aspect of an imperative or demand. Commands would thus include utterances that aren’t intended to bind, as when requests (the illocution) are put in the form of an imperative (the locution). Imagine requesting a beer from the fridge with the locution “Grab me a beer!”. This locution is imperatival but the illocution need not be. In certain contexts, for example in the context of a close friendship, it can be understood that you only intend to request the beer with this imperatival locution, not to command your friend.

On the other hand, the technical sense of ‘command’ focuses on the illocutionary aspect. Commands are utterances that are intended to bind their subjects; they are utterances with a particular illocutionary aspect. We all know that commands can come in the form of locutions that aren’t imperatives, as when a superior officer “suggests” that a subordinate act in some way. This suggestion is understood to be a command and is intended to bind the subordinate; the utterance is a suggestion in its locutionary aspect but a command in its illocutionary aspect. This utterance is a command in the technical sense but not in the everyday sense. When concerned with authority, we don’t really care what the locutionary aspect of any given command is, as long as it is clearly conveyed to the listener that the illocutionary aspect is that of command. So when I use ‘command’, I’ll be referring to utterances that are imperatives in their illocutionary aspect regardless of whether they are expressed as imperatives in their locutionary aspect.

Focusing on speech acts with a particular illocutionary intent shows that authority must be *intentional*, *performative*, and *discretionary*. First, authority is intentional.

Commands are illocutionary so require the intention of the commander to bind the subject. We know that people with authority can use imperatival locutions without commanding simply because they don't intend to bind. Imagine that your partner is the police chief in your town. She has legal authority over you and can issue commands that legally bind you. In one situation, say at a crime scene, she says "Stop!" and so legally binds you to stop: you would be acting illegally and subject to sanctions if you didn't stop. But in other circumstances, say in the privacy of your shared home, when she says "Stop!" she doesn't intend to bind you with her legal authority. The speech act is still intended to give you a reason to stop, but it isn't intended to legally bind you to stop. In both cases she uses an imperatival locution and in both cases she has genuine legal authority over you to issue these sorts of commands. The only difference is in her intention, and it is only because of her intention to bind you that you are bound. Importantly, this means subjects cannot be bound unintentionally. If your partner is in the next room and commands her deputy to "Stop!", you aren't bound merely by hearing a command: it wasn't directed at you and wasn't intended to bind you, so can't be a command that binds you.

Second, authority is performative: it must be exercised in order to bind subjects. Illocution must be conveyed by locution, so without the actual speech act, there is no command. (Note that speech acts should be understood broadly: an authority can issue written commands.) The fact that authority is performative may seem too obvious to be worth stating, but it conceals an important point. In order to be able to command, and so in order to be able to be a genuine authority, the authority must be able to convey meaning to the subject: command is an act of communication. She must be able to indicate which judgments are commands intended to bind the subject, and which other of her actions are

not intended as commands. This puts performative conditions on genuine authority. A person who otherwise fulfilled the conditions for authority but could not convey her intentions to her subjects cannot be an authority. These performative conditions are often overlooked but are an important part of authority and will be important in my account. I discuss the implications of this point more fully in later chapters.

Third, authority is discretionary: subjects are bound at the discretion of the authority. This follows from the performative nature of commands. The performance of the command is done at the discretion of the commander. I don't mean to imply that there are no standards for when a commander should command. Institutional authorities, for example military commanders, follow strict standards for how and when they command their subjects. Commands are still discretionary because the subject's standing changes at the discretion of the commander. Returning to the mother exercising her parental authority, imagine that in one case it is 7:59 when the child is commanded to go to her room. At 7:58, the child had the permission to continue playing, or to go to bed on her own, or to read, or whatever. At 8:00, though, the command has changed her situation. She now only has the permission to go to bed and the permission to continue playing has been removed by the command. In the second case the only difference is that the command is issued at 8:01. In this case, at 8:00 the child still had the permission to continue playing. These examples show that the standing of the subject is changed at the discretion of the authority. Whether continuing to play was permissible for the child at 8:00 depends entirely on the mother's discretion about when to perform the utterance and thereby issue the command.

The fact that authority is discretionary has been characterized in a variety of ways.

Raz, for example, characterizes authority as the power to bind subjects *simply by expressing an intention to do so*.¹⁰ The fact that subjects are bound simply by the authority's expression of the intention means authority is discretionary: the expression changes subjects' standing just because it was expressed, and if it hadn't been expressed, the subjects' standing wouldn't have changed. An agent acting at her discretion in this sense is acting both intentionally and performatively. So I'll characterize the commands of a genuine authority as binding subjects *at the authority's discretion*, intending to simultaneously indicate that authority is intentional, performative, and discretionary.

3. Reasons

If authority is the right to command and commands are speech acts with a specific illocutionary aspect, the next step towards understanding authority is understanding the intention that defines the illocution. When an authority issues an imperative with the intention to command her subjects, what exactly does that intention amount to? Following H.L.A. Hart and Joseph Raz, I'll take a "rationalist" approach to authority, according to which authority is best understood through the lens of the subject's reasons. When an authority commands a subject, she intends her command to constitute a particular sort of reason for the subject.

Before discussing the authority's intention in more detail, we first need to be clear about the nature of reasons, which will be a central focus of the entire dissertation.

¹⁰ See, e.g., Joseph Raz, "The Problem of Authority: Revisiting the Service Conception", *Minnesota Law Review* 90 (2006): p. 1013. [RSC]

Reasons are considerations that count in favor of.¹¹ A practical reason is a consideration in favor of acting in some way. If I am thirsty, the fact that there is water in the glass in front of me is a practical reason for me to drink from that glass. Drinking water would help quench my thirst, so the fact that there is water in the glass counts in favor of drinking from the glass. If the glass contained poison, I would have no practical reason to take a drink from the glass because drinking poison would not quench my thirst. That there is water in the glass is a reason to drink from the glass; that there is poison in the glass is not a reason to drink from the glass (assuming I aim to quench my thirst).

An epistemic reason, on the other hand, is a consideration in favor of believing some proposition. Epistemic reasons are also called evidence. If I have evidence of some proposition, I have considerations that count in favor of believing that proposition. If Xi's fingerprints are found on a knife that has been used to murder Josh, the fingerprints count in favor of believing the proposition that Xi murdered Josh. The fingerprints are evidence that Xi murdered Josh. Unlike practical reasons, evidence can be misleading. Perhaps Xi touched the knife while cutting some vegetables, but it was Helena that murdered Josh. In that case Xi's fingerprints are still evidence that she murdered Josh, they are just misleading evidence.

It is important that reasons to believe can be misleading in ways that reasons to act cannot. Perhaps the glass in front of me is filled with colorless, odorless poison that is indistinguishable from water. I have strong evidence that the glass is filled with water, so have good epistemic reason to believe that the glass is filled with water and good epistemic reason to believe that if I drink from the glass my thirst will be quenched. But because the

¹¹ I believe this is the best and fullest characterization possible, following many others who hold that the concept of a reason is basic in the sense that it can't be usefully defined in other terms.

glass is filled with poison, I do not have a practical reason to drink from the glass. I have a misleading reason to believe that I have a practical reason to drink from the glass, so I have a merely apparent practical reason.

Some theorists believe that practical reasons can be misleading in precisely the way epistemic reasons can. Another way of putting this is that these theorists and I disagree over whether practical reasons are *veridical*. I can't hope to solve the debate between these camps here. Instead, what I'll do is explain the contrasting positions and motivate why I take the veridical stance: this is intended as a frame for my project rather than as a full discussion, giving knockdown arguments against opposing positions.

Raz is a prominent example of someone who disagrees with the veridicality of practical reasons. He considers a case where an undetectable meteorite is about to crash into my town, resulting in my death.¹² Because the meteor is undetectable (not only to me, but to astronomers and everyone else as well), Raz claims I have no practical reason to leave my town. Raz and I agree that I have no reason to believe that I should leave my town because reasons to believe aren't veridical. But it seems bizarre to me to claim that I have no practical reason to leave. This is like saying I have no reason not to drink the poison if the poison is undetectable. Of course I have a reason: if I don't leave or if I drink the poison, I'll die! Survival counts heavily in favor of leaving. The fact that I don't know about the meteor is unfortunate, but it's still a reason to leave.

Part of the reason that it seems coherent and also sensible to me to hold that practical reasons are veridical is the role of third parties. Imagine that I have an omniscient

¹² Joseph Raz, *From Normativity to Responsibility* (Oxford: Oxford UP, 2011): pp. 34-5. [FNR] I'm modifying the example slightly: the meteorite is not undetectable in his case, but its movements are entirely random, so its path is undetectable.

advisor whom I can reach by phone. As a practical advisor, her role is to give advice about what I should do, so she should base her advice on my reasons. If she is half the world away and advising me about where I should eat lunch today (all on my lonesome), it would be bad advice to tell me that I shouldn't get seafood because *she* is allergic to seafood. It's only *my* allergy to seafood that is a reason for me not to have seafood. So she can only base her advice on my reasons. Returning to the meteor, if the fact that I can't know about the meteor means I have no practical reason not to leave my town, then she can't advise me to leave the town even though she knows about the meteor. This is clearly mistaken: she should tell me about the meteor and advise me to leave. Indeed, the way we choose our advisors is generally based on whether they might know more about our practical reasons than we do. Claiming that whether I have a reason to leave the town depends on whether I know about the meteor denies this.¹³

One worry about the veridical understanding of practical reasons stems from the idea that if I fail to conform to a reason I have, I am in some sense criticizable. For example,

¹³ Raz addresses something like this objection, and his reply could be modified to this case to claim that having an advisor means I *do* have a practical reason to leave the town. While I still do not know that the meteor is coming, I *can* know because I have an available, omniscient friend who does know (or can easily find out). Again, this strikes me as odd. Why should whether I have a reason to leave my town depend on somebody else's epistemic context? Here Raz's account is left intentionally underspecified. What he argues is that practical reasons are best understood as part of a "normative/explanatory nexus". The nexus includes two "filters" on reasons: a possibility filter and an epistemic filter. The possibility filter functions similarly to the "ought-implies-can" principle: if something is impossible for an agent to do, in some important sense of possibility, then it can't be the case that the agent ought to do that. The possibility filter is essentially a "reasons-imply-can" principle: even if I could save a million lives by flying around the world like Superman, I have no reason to fly around the world like Superman because I can't. The epistemic filter does the same work, but the impossibility is epistemic: if an agent can't know about some fact, that fact can't provide a reason for the agent. The fact that I can't know about the meteorite means its impending impact gives me no reason to leave my town. But Raz admits that the sense of possibility in the epistemic filter is unclear. He wants to separate cases like the meteorite, where there is absolutely no plausible way I could know that the meteorite is about to hit my town, from cases of "superficial epistemic incapacity" (FNR, p. 126.) He thinks it is clear that epistemic incapacity sometimes affects whether we have practical reasons, but it is equally clear that epistemic incapacity sometimes *doesn't* affect whether we have practical reasons. So he attempts to capture both points while leaving the dividing line between them vague. I don't have the space to get into the details of this interesting account. As noted, at this point I'm simply trying to elucidate the veridical position and give some reason why I prefer it without pretending to have decided the issues.

perhaps the fact that I failed to conform to my practical reason to leave my town means that I am *irrational*. After all, irrationality is centrally concerned with responsiveness to reasons. If the charge of irrationality is based in my conformity to my practical reasons and I have a practical reason to leave my town, then I am irrational for failing to leave even though the meteor was completely undetectable. It is clear that the charge of irrationality is misplaced in this case. But that's not because I lack a practical reason to leave my town. It's because the charge of irrationality is not based in conformity to practical reasons but conformity to epistemic reasons. If I saw the meteor heading towards my town and failed to leave, that would be irrational. There is much more to be said about the choice between veridical and non-veridical accounts of practical reasons, and I don't expect that opponents will be convinced by this discussion. But from here on I will assume that practical reasons are veridical (but that epistemic reasons are not).

I'll distinguish among reasons by grouping them into *domains*. Morality, prudence, basketball, physics, and anthropology are examples of domains of reasons. A classic moral reason is whether some act inflicts unnecessary pain on an innocent. The fact that an act inflicts unnecessary pain on an innocent is a moral consideration in favor of not performing that act. Similarly, the fact that an act would be very harmful to me is a prudential consideration in favor of me not performing that act. The domains of prudence and morality are at least in part defined by their aims. Prudence, for example, aims at maximizing personal welfare (or something of the sort). That an act would be harmful to me is a prudential reason not to perform that act because harm does not serve the aim of personal welfare. Basketball is an example of a practice-based domain. Basketball is a social practice with a set of defined aims (winning games, by scoring the most points, or

winning tournaments, by winning games) and rules. If I have the ball and am undefended under the basket, I have good basketball reason to shoot the ball: shooting the ball serves the aim of winning the game. If I am a professional basketball player and my salary depends on my performance, I have both a basketball reason and a prudential reason to shoot the ball.

Distinguishing reasons in this way is useful but quite complex. Some domains overlap, as prudence does with both basketball and morality (the fact that an act would harm me is both a prudential and a moral reason not to perform it, although that reason may not tip the balance of reasons in either domain). Further, as practical agents we try to act on our all-things-considered balance of reasons. This means that we have to balance reasons from different domains against each other. For example, a relatively common understanding of morality holds that moral reasons are lexically ordered with respect to reasons from all other domains. This would mean that if the balance of moral reasons favored one option, it doesn't matter what the balance of reasons from any other domain favors. Your all-things-considered balance of reasons is to do what the balance of moral reasons favors because moral reasons tip the scales in the all-things-considered balancing. Most domains, perhaps all, aren't lexically ordered and so balancing them against one another is not easy. Determining what act is supported by the all-things-considered balance of reasons will often be massively complicated, requiring one to balance reasons from different domains (and reasons that count differently in different domains). The complex nature of grouping reasons by domain is unavoidable. It simply reflects the fact that the world is very complex so acting well in the world is no easy affair.

Attempting to come to an all-things-considered balance of reasons is difficult

because we often, perhaps always, have reasons that conflict. Practical reasons conflict when they count in favor of doing conflicting actions, e.g. when performing one act entails not performing another, while epistemic reasons conflict when they count in favor of believing conflicting propositions. When reasons come into conflict, one can *defeat* another (or a group of reasons can defeat another group). There are different kinds of defeat. The most common and obvious kind of defeat is usually put in terms of outweighing, although the metaphor of force is also used. Let's say you're considering going to the gym. The fact that you will get exercise counts in favor of going to the gym, while the fact that you will be sore afterward counts in favor of not going to the gym. Most days, the exercise is a greater gain than the soreness is a loss, so exercise outweighs soreness and the balance of reasons favors going to the gym.

Crucially, when a reason is outweighed, it still retains its force. The soreness is still bad, and that badness remains even though, on balance, it's better to accept the soreness in order to also get the benefits of going to the gym. The fact that the soreness is merely outweighed, though, means I have reason to take action to mitigate my soreness even when I ultimately decide to go to the gym, for example by stretching beforehand. This idea is perhaps clearest in the moral case. If I have made a promise to meet my friend for lunch but on the way I encounter a drowning child that I can easily rescue, the reasons I have to save the child outweigh the reasons I have to keep my promise and meet my friend. The fact that this reason is merely outweighed means I have reasons to compensate for my failure to keep my promise; for example, I should apologize and perhaps set up another time for lunch (although the fact that my promise was defeated means I did nothing wrong by breaking it). In cases of outweighing there is a "remainder", because the outweighed

reasons are still reasons with force that must be accounted for.

Another kind of defeat is canceling. In contrast to outweighing, when a reason is canceled there is no remainder. Perhaps I need to get to the airport as fast as possible. Of my current options, taking the metro will be quickest, so I have reason to take the metro. However, just as I am about to embark, my friend tells me she can drive me to the airport, and this would be significantly faster than taking the metro. I now have no reason to take the metro; the reason I had was canceled by the possibility of getting a ride. There is no remainder and I should feel no regret about not taking the metro. But I would still have reason to take the metro if the car weren't an option.

This fact contrasts canceling with undercutting. A reason is undercut when the prima facie connection between the fact and what it counts in favor of is severed. Here's an example modified from a classic case in the epistemology literature. I am looking at what appears to be a barn: wooden, painted red, large, tin roof, and so on. My perceptions give me good reason to believe I am looking at a barn. But I find out that I am in a county where most of the barn-looking buildings are actually clever movie set mock-ups. This fact undercuts the normally reliable connection between my perceptions and what is the case. Breaking this connection undercuts any reason I may have had to believe this is a barn. A canceled reason can be rehabilitated by changing other reasons; for example if my friend rescinds her offer to drive me to the airport, I now have reason to take the metro again. But an undercut reason loses its status as a reason more permanently because the connection between the fact, like my perception of barn-like characteristics, and what it counts in favor of, like whether there is a barn there, is severed.

4. Commands as Content-Independent Reasons

With this understanding of reasons we can more fully characterize authority. Following the rationalist approach, authorities are distinguished by the kind of reasons their commands constitute for their subjects. In this way we can distinguish both authority from other kinds of speech acts and types of authority from each other.

At the end of section two, we came to the conclusion that commands were speech acts with a particular illocutionary aspect. The authority intends her command to be a particular kind of reason for subjects, namely a *content-independent* and *preemptive* reason. Both features contrast authority with related but distinct speech acts, and neither is a particularly worrying feature of a reason in general. In combination, though, they make clear what is distinctive about authority and why genuine authority is difficult to attain. I address content-independence in this section and preemption in the next.

In general, an utterance has two important features that affect what kind of reason it constitutes. These are the utterance's content and provenance. Commands are interesting because they constitute *content-independent* reasons for subjects. Raz and Hart develop the idea of content-independence from Hobbes' characterization of command. Hobbes says, "COMMAND is, where a man saith, *Doe this*, or *Doe not this*, without expecting any other reason than the Will of him that sayes it."¹⁴ According to Hobbes, the commander can command opposites (do this or do not this) but bind the subject in either case. A first pass at the idea of content-independence is that an authority's command binds regardless of content. The content is irrelevant to whether the subject has a reason to do as commanded.

To understand content-independence a useful comparison is threats. Imagine you

¹⁴ Thomas Hobbes, ed. Richard Tuck, *Leviathan* (Cambridge: Cambridge UP, 1991), Chapter 25, p. 176.

are kidnapped and threatened with your death and the death of your family unless you... well, it doesn't matter what demands are made on you, and that's the whole point. The reason to comply is the same and has the same strength regardless of whether your kidnapper demands money, or information, or an act. The reason to do as you are threatened is independent of the content of the demand.

If the content is irrelevant to whether the command constitutes a reason for subjects, then the reason must come from the command's provenance instead. As Hobbes says, the reason to act as commanded is not whether you are commanded to do a particular thing but because of "the Will of him that sayes it". Authorities claim they should be obeyed because *they* issued the command. Imagine a parent orders her child to go to her room. If the child asks why, a sufficient response would be "Because *I* told you to!". Appealing to the provenance of the command as the reason to obey is constitutive of authority. If the parent had responded, "Because if you don't, you're grounded!" or "Because if you do, I'll pay you.", she wouldn't be appealing to her authority. In the first case the reason to obey is the threat and in the second case the reason to obey is the benefits the child will gain from obeying.¹⁵

Had a stranger issued an identical order for the child to go to her room, the child would have no reason to obey because the stranger is not the right sort of person to issue binding commands to the child: the command's provenance is wrong. Imagine you are walking down the sidewalk and an approaching stranger says "I command you not to walk there!". You may respond, paradigmatically, "And just who do you think you are?". This is a good response because you are challenging the authority of the stranger: you are questioning whether the stranger is the sort of person who can bind you with commands,

¹⁵ The threat may be an indirect reference to authority, if the claim assumes that the grounding is justified because the child disobeyed a command from a genuine authority.

i.e. you are questioning the command's provenance.

So the reason constituted by a command is provenance-dependent. To say the reason is provenance-dependent is just to say that altering certain characteristics of the provenance of the command can alter the strength of the reason, or even whether there is such a reason at all. Provenance-dependence is not a unique feature to authority, and in fact is ubiquitous. If I ask a financial advisor and my twelve-year old which stocks to invest in, I ought to weight the advisor's answer much more heavily than my child's. The financial provenance of one judgment is much better than the other.

While provenance-dependence is a common feature of reasons, what's uncommon is *complete* provenance-dependence. This is equivalent to content-independence on the assumption that provenance and content are the only relevant features of utterances. If content is ruled out as a source of the reason to obey, then the only remaining source is provenance. As Richard Friedman puts it,

What is therefore essential to the concept of an authoritative command is the opening up of a distinction between the person who prescribes and what he prescribes, so that the content of the prescription becomes irrelevant, and the person becomes the fact that endows the prescription with its distinctive appeal.¹⁶

This is odd. Normally the reasons that justify my act are directly related to the act. I go to the fridge because the food is there, I do pushups because of the health benefits, and so on. There is a direct connection between the reason (health benefits) and the act (pushups). But content-independent reasons sever this direct connection. If I do pushups because my superior officer commanded me to, it's not because of anything about pushups. I could have been commanded to do sit-ups, or to stop doing pushups I was already doing, and I

¹⁶ Richard B. Friedman, "On the Concept of Authority in Political Philosophy", in ed. Raz, *Authority*, p. 66.

would have the same reason to act. When acting because of a command, the connection between the reason the subject is acting and the act itself is thus indirect.¹⁷

There is one important qualification and one important implication of this discussion. The qualification is that the first pass at content-independence was stated too strongly. The first pass was that the content of a command is completely irrelevant for whether the command constitutes a reason for subjects. But the content of a command is not *completely* irrelevant. The most obvious exception is cases where subjects are commanded to do something clearly unjust. If my military commander directs me to commit genocide, that doesn't mean I must commit genocide in order to conform to my reasons. Instead I have a reason to believe that the commander lacks authority over me in this domain. Egregious errors undermine authority.

This is analogous to limitations on two other content-independent speech acts. The first is threats, mentioned above. There I said that if you are threatened with the death of your family unless you act in some way, you have a reason to act in the way directed no matter what that act is: paying a ransom, stealing work product, and so on. But this isn't quite right because at some point, the threatened act can be so egregious that you shouldn't do it even though your family will die if you don't. If the threat is that your family dies unless you set off a nuclear weapon in New York City, the content of the threat matters to whether you have a reason to comply.

The other relevant comparison is to promises. Promises give promisers content-

¹⁷ Ibid. See also Raz, MF, p. 35. He also calls the command "extraneous" to the act.

independent reasons to do as promised.¹⁸ If I promise to meet you at Myrtle's for lunch tomorrow, I have a reason to go to Myrtle's. Had I promised to meet you at Joe's instead, I would have an identical reason to go to Joe's. The reason is the promise, not the particular content. But there are some things I can't promise to do, and even if I make a "promise" to do them, that "promise" doesn't give me a reason to act. Murder is an example: I can't promise to murder someone and then claim my killing was justified because I made a promise. The promise doesn't bind me because of its content, even though in general the reason to do as you promised is content-independent.

The idea that commands, promises, and threats constitute content-independent reasons except in some cases smacks of incoherence and perhaps looks ad hoc. Why does content matter in some cases but not others? And if content is sometimes relevant, why does it make sense to call these content-independent reasons? Threats are a degenerate case that I'm going to put aside. In the cases of promising and commanding we have a good explanation for this seeming incoherence. If I say something like "I promise to murder Joe", I have made a failed attempt at promising. It's not that I promised to murder Joe and then my promise was somehow defeated by the content of the promise. The thing I tried to promise to do I simply cannot promise to do. The content of the utterance meant it couldn't be a promise.

The same is true of commands. When clearly unjust commands are issued, it's not that the subject is bound by the command and then the content defeats the command. It's that the commander never had the authority to command that in the first place. She's

¹⁸ Many make this comparison between commands and promises, including Hobbes, Hart, Raz, and John Rawls, "Two Concepts of Rules", *The Philosophical Review* 64 (1955): pp. 13ff. Hurd is one of the few examples of someone who denies that promises constitute content-independent reasons.

issuing a command that falls outside her domain of authority, so it doesn't bind. It's like your boss trying to command you how to raise your kids. Her authority doesn't extend to your parenting, so when she issues "commands" about your parenting, you aren't bound. Of course it can be difficult to determine when an authority is issuing commands outside her domain. It's pretty clear in the boss case because parenting and occupation are rather distinct domains. It's less clear in a case like a military commander ordering genocide because military commanders generally have authority in domains like killing. (Specifying the precise boundaries of an authority's domain is thus a very important task.)

Content-independence is best specified as the claim that the content of a directive that falls within the authority's domain is irrelevant. When the directive falls outside the authority's domain, the content can be a good indicator that the directive doesn't bind. But the reason it doesn't bind is because it fell outside the authority's domain, not that the content was so mistaken (as determined by comparison with reasons within the authority's domain) that it was defeated by other reasons.

Importantly, content-independence entails that *some mistaken commands bind*. I'll clarify the idea of mistaken commands below, but the essential point is that the balance of reasons in a particular domain will favor some act independently of the command. When an authority's directives demand that subjects act in a way that is contrary to how the balance of reasons in that domain would otherwise dictate, the command is mistaken. Mistakes, then, are an issue of content: whether the content matches to the preexisting balance of reasons in the domain. But if the reason to obey is content-independent, then whether a command is mistaken can't affect whether the command constitutes a reason for subjects.

If mistaken commands didn't bind, then authority wouldn't give content-independent reasons. Instead, every time their directives were mistaken they would fail to bind. In the end this simply entails that authority never changes people's reasons: either the directive is accurate and tells them what their reasons entail anyway, or the directive is mistaken so doesn't bind. But changing reasons is exactly what authority does. As Raz bluntly puts it, "there is no point in having authorities unless their determinations are binding even if mistaken".¹⁹ If authority's commands only bound subjects when they were correct or just, then authority becomes epiphenomenal.

Another way of putting this is that without the possibility of mistaken but binding commands, authority is just advice. But there are clear cases where authority is more than advice, like when you consent to obey another person, and our practices of authority make it clear that it is supposed to be different in kind than advice. Content-independence is a necessary and defining feature of authority. But it is only in combination with the fact that the commands of a genuine authority also constitute preemptive reasons for subjects that content-independence becomes worrying.

5. Commands as Preemptive Reasons

The commands of genuine authorities constitute preemptive reasons for their subjects. Preemption allows us to capture the sense in which subjects are bound by commands. (In chapter three I discuss this in considerable detail.) When a practical command is issued, it claims to be a reason to act in some way. But in many cases there are obvious and weighty reasons to act contrary to the command. This is especially true when the command is

¹⁹ Raz, MF, p. 47.

mistaken, as defined above. When a mistaken command is issued, it counts in favor of doing as commanded. But if there are reasons from within the domain of the authority that count in favor of not doing as commanded, why can't subjects simply ignore the command and conform to the other reasons?

There are two worrying implications if we can't answer this question. First, if subjects can simply disregard mistaken commands and do as the reasons otherwise indicate, authority is again rendered useless. It's the same problem as the case where the commands of an authority are not content-independent. In this case mistaken commands are still commands but they are simply overridden by the other, countervailing reasons. Second, if subjects can conform better to their reasons by disobeying an authority, it looks irrational to obey authorities. This is one of the central problems theories of authority have to grapple with.²⁰ Preemption explains how subjects are decisively bound by commands, so ultimately explains how subjects are rational to obey authorities.

To understand preemption, a useful comparison is with an arbitrator.²¹ When conflicting parties submit their problem to an arbitrator, they agree to follow the judgment of the arbitrator even if she rules against them. They may think the balance of reasons is in their favor, for example justifying recompense. But when the arbitrator issues her judgment, the parties' personal judgments about the balance of reasons and the final decision have to be put aside. Her judgment is decisive because it tilts the parties' balance of reasons inexorably in its own favor: in order to act on their reasons, they now must put aside their own judgment and act as the arbitrator decided.

²⁰ For the most trenchant expressions of this concern, see Wolff and Hurd, but it is also discussed by Raz, Friedman, and many others.

²¹ Cf. Raz, MF, pp. 41-2.

Arbitrators and authorities contrast in this respect with advice.²² When you get advice, it is merely one part of the overall balance of reasons. You may decide to accept the advice or reject it, but it is not decisive. Your balance of reasons may still favor going against the advice. Arbitrators and authorities claim to do much more. Once their judgment is made, you cannot simply factor it into your judgment of the balance of reasons and proceed as usual.

Arbitrators' judgments and the commands of genuine authorities are decisive because they are preemptive. As Raz says, "the fact that an authority requires performance of an action is a reason for its performance which is not to be added to all other relevant reasons when assessing what to do. It should *exclude and take the place* of some of them."²³ Preemptive reasons are decisive because they defeat countervailing reasons in a way *merely additive* reasons, like that given by advice, do not. Above I noted that conflicting reasons are subject to three kinds of defeat: outweighing, canceling, and undercutting. But in what way do the preemptive reasons given by authorities' commands defeat countervailing reasons? This is a surprisingly thorny question to answer. Here I'm going to sketch a preliminary answer that is to my mind largely correct, but will considerably refine the answer in chapter three.

On Raz's account, following Hart, the commands of genuine authorities constitute a two-part reason, variously called a preemptive, peremptory, or protected reason.²⁴ A preemptive reason consists in two reasons: a simple first-order reason that counts in favor

²² Hobbes was perhaps the first to contrast command with advice, a theme picked up by Hart and Raz, among others.

²³ Raz, MF, p. 46. My emphasis. Raz calls this the preemption thesis.

²⁴ The notion of a preemptive reason is central to Raz's account of authority, but also to his account of rules and norms in general. See *Practical Reasons and Norms* (Princeton: Princeton UP, 1975) [PRN] and *The Authority of Law* (Oxford: Oxford UP, 1979): p. 27. [AL]

of doing as commanded and a second-order reason. The force of the first-order reason is something akin to a request. The more important part of preemptive reasons is the second-order part, what Raz calls an *exclusionary* reason.

Exclusion is Raz's novel notion of second-order defeat. When reasons conflict in the first-order, they are subject to the normal kinds of defeat. Sometimes, though, we have reasons to treat our reasons in certain ways: second-order reasons. Exclusionary reasons are considerations that count in favor of not acting on other reasons. Importantly, exclusion and other kinds of defeat are distinct. At the first-order level, some reason could be hugely weighty yet still be excluded by a second-order reason. The weight doesn't affect whether it is excluded. Raz argues that the commands of a genuine authority exclude countervailing reasons and this explains why preemptive reasons are decisive.

In the case of practical authority, a command gives an exclusionary reason not to act on reasons that count in favor of acting in any way other than as commanded. Reasons that would support acting in any other way are excluded from the subject's practical deliberation.²⁵ To illustrate this, imagine an authority has just commanded you to go to the gym (perhaps you promised to obey your partner's gym-directives in an effort to take the control out of your akratic hands). You already had some preexisting reasons to go to the gym, most obviously the health benefits. You also had some preexisting reasons not to go to the gym, like the pain in your knee and the desire to get some lunch.

Now the authority commands you to go to the gym, and you have a new reason to go to the gym: you were commanded to. The command also constitutes a second-order reason

²⁵ Here I characterize Raz's most recent position with respect to exclusionary reasons; in earlier writing like PRN he claimed *all* reasons other than the fact of the command, not only the countervailing reasons, are excluded. This nuance is discussed in Appendix A.

to exclude any reason that counts in favor of not going to they gym. Before you may have judged that avoiding more pain in your knee and getting lunch would outweigh going to the gym, and so not gone to the gym. But the pain in your knee and getting lunch are ruled out as reasons for you to act by the exclusionary reason. The effect of this is that, in order to conform to the balance of your reasons, you must go to the gym. This is the only act that has any support from reasons; the reasons that count in favor of not going to the gym have been excluded.

A common misunderstanding of authority follows from misunderstanding preemption. It is tempting to assume that preemptive reasons must be all-things-considered or absolutely decisive; after all, countervailing reasons have been excluded. Another way of putting this is that it is attempting to assume that preemptive reasons must be undefeatable. If the commands of a genuine authority are absolutely decisive, disobeying such commands is necessarily wrong. You would have to obey come hell or high water. If this were the case, we would be justified in rejecting the notion of authority altogether. So to forestall this misunderstanding and accompanying objections, we need to be very precise about what preemption entails.

Authorities' commands are not absolutely decisive because they only preempt certain reasons. Preemption is constrained by three factors: the authority's jurisdiction, the available preexisting reasons, and the moment of performance.²⁶ First, jurisdiction. Above I defined an authority's jurisdiction as the domain or domains it has authority over as well as the agent or agents it has authority over. Your boss' jurisdiction is a single domain and over multiple individuals. She has occupational authority over you and the

²⁶ Cf. Raz, MF, pp. 46, 62.

other employees; she lacks authority over any employee with respect to non-occupational domains and she lacks authority over any non-employee with respect to all domains. If she tries to issue commands that fall outside this jurisdiction in either respect, her directives fail to preempt and so can be defeated. So if she tries to issue a command to you about how you raise your children, this directive does not decisively change the balance of your parenting reasons because her authority does not extend to the parenting domain. Similarly, if she tries to command any non-employee, the balance of their reasons does not decisively change because her authority does not extend to them and doesn't preempt any of their reasons. In later chapters I'll argue that systematically issuing commands outside jurisdiction is a defeater on authority.

Further, even when her commands fall within her domain, they are only preemptive in that domain. If she orders you to work this weekend, the balance of your occupational reasons decisively changes such that you should now work this weekend. But if your partner becomes seriously sick and is rushed to the hospital, your all-things-considered balance of reasons decisively favors leaving work and going to the hospital. Your boss's command was still preemptive within the occupational domain, but it is not preemptive outside the domain of her jurisdiction. Your boss's occupational authority can't exclude reasons in other domains.

Second, authorities' commands only constitute preemptive reasons with respect to those reasons within their domain that are accessible to them at the moment of command. The notion of exclusion clarifies this constraint. When an authority issues a command, it excludes countervailing reasons within her jurisdiction. But it doesn't exclude all possible countervailing reasons within her jurisdiction, it only excludes the ones that she has

available to her. Again, there is an illuminating parallel with arbitrators. Imagine that you and an ex-partner submit to arbitration on the issue of parental rights. You know that your ex-partner is an addict who often ignores the welfare of your child in favor of feeding the addiction. Your ex successfully conceals the addiction from the arbitrator, who issues a judgment rewarding your ex full custody. This judgment doesn't decisively bind you because in this case you know that there is a reason from within the relevant domain that was unavailable to the arbitrator. Now, the judgment may still give you instrumental reasons to hand over custody, or reasons to reexamine your own parental fitness, but the fact that the arbitrator didn't know about the debilitating addiction means that the arbitrator's judgment isn't decisive for you. Of course it is difficult to judge when the arbitrator actually lacked certain information as opposed to when you merely hope she lacked certain information because she ruled against you. Again, though, this pragmatic difficulty is no barrier to the conceptual point that if certain reasons from within the authority's jurisdictional domain are unavailable to her, then her directives might not decisively change the balance of your reasons, even in that domain.

Third, authority's commands only constitute preemptive reasons at the moment of command. New reasons from the very same domain as the authority's jurisdiction can render the authority's command inert. If we just modify the arbitration example, imagine that at the time of arbitration your ex is not an addict, but in the months afterwards develops an addiction that puts your child at imminent risk of severe harm. The addiction was unforeseeable, but once it arises it overrides the arbitrator's directive. Again, practical matters are more complicated: the fact that you might be (unjustifiably) punished if you don't conform to the arbitrator's judgment gives you a reason to follow it even though it

doesn't bind you anymore, and perhaps the best response to the lack of bindingness is to ask the arbitrator to decide anew rather than just taking your child back. But if your ex's addiction is putting your child in imminent risk of severe harm, it might be that your only option is to take your child. These pragmatic points don't undermine the fact that new reasons can arise even from within the authority's jurisdiction. As a technical matter this is simply an extension of the second constraint on preemption, but it's an important and noteworthy extension. Many reasons that arise after the command are unlikely to have been available to the authority at the time of command, although certainly not all of them.

In sum: the commands of a genuine authority constitute preemptive reasons within their jurisdictional domain for subjects. The preemptive effect of the command excludes some reasons from the balance of reasons within that domain and so changes the balance of reasons for subjects in a non-additive way. Only those reasons that fall within the authority's jurisdictional domain and are available to her at the time of command are preempted. Reasons that aren't preempted are still part of subjects' all-things-considered balance of reasons, so sometimes (or even often!) the subject won't be required to obey the authority's command in order to fully conform to her reasons all-things-considered.

The fact that the subject sometimes doesn't have to obey the authority's command doesn't mean that the authority wasn't genuine or that the command failed to bind. It did bind and so did have a preemptive effect on the subject's balance of reasons, it's just that the effect was constrained. We tend to think of authority in its most extreme forms, like god exercising divine moral authority over people that binds absolutely. If authority must be absolute in this way, then the fact that subjects often disobey yet still act on the balance of their reasons makes it appear that the command didn't bind. But once we see that

preemption is constrained, the pragmatic disconnect between our practices of obeying authorities of various kinds and the theoretical picture of authority as preemptive is not worrying.

6. Authority Changes Standing

Up to this point I have been intentionally coy about one of the central aspects of authority. So far, we have seen that authorities issue commands, discretionary speech acts that constitute preemptive, content-independent reasons for subjects. But what does a preemptive, content-independent reason *do*? In what way should we characterize the effect of the commands of a genuine authority on a subject's balance of reasons?

Our starting point was traditional authority, understood as the *right* to command that creates a new *duty* for subjects. According to this account what preemptive, content-independent reasons do is create duties; subjects are bound because they have a new duty. The thought is that when the balance of reasons decisively favors some act, I have a duty to perform that act. Thus, preemptive reasons create duties: they exclude all countervailing reasons and add a new reason in favor of doing as commanded. Exclusion rigs the balance of reasons in favor of doing as commanded, so a command tilts the balance and creates a duty to do as commanded. This understanding of authority, however, is too narrow. There is an alternative, and very sensible, way to think of authority. Authority is not a right. Authority is a *power*.²⁷

²⁷ This point is indebted to Arthur Isak Applbaum, "Legitimacy without the Duty to Obey", *Philosophy and Public Affairs* 38 (2010): pp. 215-239. See also David Copp, "The Idea of Political Legitimacy", *Philosophy and Public Affairs* 28 (1999): pp. 3-45.

The contrast between right and power is found in Wesley Hohfeld's framework.²⁸ Hohfeld looked at the way 'right' was variously used in legal contexts and distinguished four correlated pairs of legal advantage and disadvantage, the sum of which comprises a person's legal standing. We can take the notion of legal advantage and disadvantage and apply it to morality and other normative domains.²⁹ For example, a person's moral standing is a complete description of her moral advantages and disadvantages.

The first, paradigmatic pair of advantage and disadvantage is claim-right/duty. If A has a claim-right against B, then B has a correlative duty to A. A's advantage is the claim-right, B's disadvantage the duty. The second correlated pair is vulnerability/liberty: if A has a liberty with respect to B, B is correlatively vulnerable to A.³⁰ To see the contrast between the pairs, imagine there is an apple between us. If I bought the apple, I have a claim-right to it as my property, so you have a correlative duty not to take it and wrong me if you do. On the other hand, if the apple fell off a nearby, unowned apple tree then I merely have a liberty with respect to it, so am vulnerable to your taking it and you do not wrong me if you do. You also have a liberty to the unowned apple so are vulnerable to my taking it and I do not wrong you if I do.

Traditional authority is the claim-right to command. Because it is a claim-right, it entails a correlative duty. Authorities have the right to command subjects, so subjects have a correlative duty to obey the command. The connection between the nature of authority and how subjects are bound is direct and intuitive. But this notion of authority cannot

²⁸ Wesley Newcomb Hohfeld, *Fundamental Legal Conceptions as Applied to Judicial Reasoning* (New Haven: Yale UP, 1919).

²⁹ See Applbaum, "Legitimacy without the Duty to Obey", pp. 220ff.

³⁰ The terminology for this pair varies. "Vulnerability" is my own amendment to Hohfeld's scheme. Hohfeld uses "no-claim" or "no-right". What I call "liberty" is also sometimes called "permission", "privilege" and "justification-right".

capture some paradigmatic exercises of authority.

Consider a police officer exercising her legal authority over you. Legal authority is a form of practical authority, changing the legal reasons you have to act in some way. When a police officer commands you to stop, you gain a legal duty to stop and so act illegally if you fail to conform to that reason and don't stop. From the legal perspective, the command of a police officer to stop is a preemptive reason, excluding other legal reasons. This is clearly an exercise of authority, imposing a legal duty.

But then the officer tells you that you may continue. This utterance cancels your legal duty to stop. You are now at legal liberty to continue, but you are also at legal liberty to remain stopped (assuming the officer didn't also command you to go on). When the officer tells you that you may continue, she is exercising her legal authority over you, but she can't be exercising traditional authority. Her utterance doesn't give you new legal duties and you don't have any new legal reason to do or not do anything. Instead, you lose your legal reason to stop and you gain legal liberty.

Despite the fact that you don't gain a new legal duty, her utterance was a content-independent and preemptive legal reason for you. Imagine that she commanded you to stop because of a chemical spill and it was too dangerous for you to continue. She's following regulations and knows that a certain risk of exposure means she should cordon off civilians, while a risk of exposure lower than that threshold means she shouldn't impede traffic and should allow people through. In this case she misjudges the risk of exposure and thinks that it has gone below the threshold such that she should allow people to continue on. If her utterance was a content-*dependent* legal reason for you, then you would still have a legal duty to stop. After all, she was mistaken about how dangerous it is. But if you

continue on after being given mistaken permission, you do not act illegally. This is because your legal duty to stop was canceled by the content-independent reason given by the later utterance.

Her utterance also gives you a preemptive reason because it excluded other legal reasons you may have had to stop. In fact, it canceled your earlier duty. This is like a police officer in an intersection waving you through a red light. You had a legal reason to stop, but the police officer's command excluded that reason and gave you a liberty to go through the red. My point is this: the only way to accurately capture the effect that the police officer's second utterance had on you is to see that this utterance also gave you a content-independent and preemptive reason. Her latter duty-canceling, liberty-making utterance of the police officer was as much an exercise of her authority as her former duty-making, liberty-canceling utterance. This shows that the notion of authority as only duty-making must be mistaken. Otherwise the latter utterance would not have been an exercise of the officer's legal authority. This ignores both the fact that the utterance superficially appears to be an exercise of authority and the fact that the utterance constituted a content-independent, preemptive reason for you.

Authority must, conceptually, give a content-independent and preemptive reason that changes your standing, but it doesn't matter what part of your standing is decisively changed. Above I defined an agent's normative standing as the sum of her advantages and disadvantages within that normative domain, so your legal standing is the sum of your legal advantages and disadvantages. In other words, it is the complete specification of your legal duties, legal rights, legal vulnerabilities, and legal liberties at a particular time (as well as your second-order legal advantages and disadvantages, which I discuss shortly). So

authority can be duty-making as the traditional account claims, but it can also be duty-canceling, liberty-making, vulnerability-making, vulnerability-canceling, and so forth. A has authority over B in domain D iff A's directives constitute a preemptive, content-independent D-reason for B and so necessarily change B's D-standing.

One way to be misled about the fact that authority can be exercised by changing parts of standing without changing duties is to assume that the commands of an authority must be preemptive reasons (which is true) and that preemptive reasons must be duty-making (which isn't). To see why this tempting thought is false, consider that if the apple really is unowned then that fact is a preemptive reason that entails both of us *lack* a duty not to take it. That is, the facts about the lack of ownership are decisively liberty-making and vulnerability-making: the fact that neither of us owns it means both of us are at liberty to take it and both of us are vulnerable to it being taken by the other. These reasons are decisive because they create the liberty regardless of other facts; even if I have more reason to take the apple than you do, for example if apples are my favorite food and your least favorite food, I do not thereby gain a moral right to the apple. You are still at liberty to take it and do not wrong me if you do because the facts about the lack of ownership are decisive liberty-making reasons.

This characterization makes it clear that authority is a power, not a right. Power is a second-order advantage. It's second-order because it references first-order advantages, and in particular a power allows one agent to change another's first-order standing. Like at the first level, Hohfeld identifies two second-order connected pairs of advantage and disadvantage. They are power/liability and immunity/disability.

If A has a power over B, B is liable to A. If I let you borrow my book, I have the

power to give you a new duty to return the book to me. I may not have specified a time, but if I say “I need the book back by tomorrow” you have a new duty to return it to me by tomorrow. You were liable to a change in your first-order normative standing, a new duty. My power can also change your powers, as when a sheriff exercises her legal power to deputize someone and grant them various legal powers. The other second-order pair is immunity/disability. If A is immune to B, B has a disability to A. You and I are legally immune to having our innocence or guilt for a crime determined by each other, and we are similarly disabled against each other in that respect (assuming one of us is not on a jury or a judge).

The police officer exercises authority over you, but with this schema it is clear that she is exercising a power over you, not exercising her rights. When she issued her command to stop, you gained a new legal duty. When she said you could continue, you lost a duty and gained a liberty. This is precisely what having a normative power over another agent allows.³¹ If authority is a right, it can only create new (correlative) duties and can't create new liberties. To capture the officer's legal authority, we must conceive of authority as a power. Instead of thinking of authority as the *right to command*, it is more apt to consider it the *power to bind*.

With the concept of authority as a normative power in hand, new possibilities have opened. In particular, if authority is a power then it is possible for authorities to change subjects' standing in ways other than by the creation of new duties. In the next chapter I'll argue that this possibility is important because authority that creates new duties is unnecessary for political authority.

³¹ Note that Raz also uses the term 'normative power' but means something quite different than I do; he's referring to a more general idea than a Hohfeldian advantage. See, e.g., MF, p. 24.

7. The Service Conception of Authority

To this point I have been discussing the concept of authority and have argued that the concept of traditional authority is too narrow. Instead of simply creating duties, authority is a normative power that can change any component of subjects' first-order normative standing. At her discretion, an authority can issue a directive that constitutes a preemptive, content-independent reason that binds subjects by changing their standing. With this concept in mind, we can now turn to a conception of genuine authority.³² Here we are not concerned with what authority is, but with how an agent can attain genuine authority. In short, we are looking for the necessary and sufficient conditions on genuine authority.

If our concept of authority is accurate, then the conditions on genuine authority will be the same as the conditions on one agent's discretionary utterances constituting preemptive and content-independent reasons for another agent in a particular domain. One way for utterances to do this is via consent. I can consent to treat your utterances as preemptive and content-independent reasons and thereby make you an authority over me. But often in discussions of authority we are concerned with whether one person could have authority over another without their consent, especially with respect to political authority. I'll explain these ideas more fully in the next chapter. For now the important idea is that there is no single set of necessary and sufficient conditions on genuine authority. Consent is one way to get genuine authority, but there are others, and these others will be quite different.

The most widely accepted and influential conception of authority is Joseph Raz's

³² John Rawls, *A Theory of Justice*, Revised Edition (Cambridge: Harvard UP, 1999): p. 5.

service conception.³³ I've already followed Raz's approach in a variety of ways, particularly in describing the commands of authority as constituting preemptive and content-independent reasons. In general I think the service conception is basically correct, even once we have expanded our concept of authority to include any change in standing rather than just focusing on duties. In following chapters I'll offer some arguments against details of Raz's view and so will modify the service conception.

The service conception is motivated in large part by the rationality challenge to authority. In connection with preemption, I noted that subjects look irrational if they obey commands, especially mistaken commands, because they are not conforming to their reasons. It is only because reasons for acting in ways not commanded are preempted that subjects could be acting on the balance of their reasons by obeying an authority. The first important condition on genuine authority comes from a similar line of thought.

Above I briefly argued that good advice must be based on reasons that already apply to a subject. If my lunch advisor tells me not to get seafood because she is allergic to seafood, this is bad advice because it takes her reasons to be my reasons. I am not allergic, so I have no reason to avoid seafood. In order to explain how obeying an authority could be rational, it must be the case that genuine authority is the same way. If an authority were commanding based on her reasons for action rather than on yours, her commands give you no reason to obey. Instead, binding commands must be based in your reasons. Raz calls

³³ Raz's theory is explicated over many years, books, and articles. I will mainly be working from MF, and I have already cited AL, PRN, RSC, and FNR. To complete the picture, I will also use "Facing Up: A Reply", *Southern California Law Review* 62 (1988): pp. 1153-1235 [FU] and *Between Authority and Interpretation: On the Theory of Law and Practical Reason* (Oxford: Oxford UP, 2010) [BAI]. For a recent, concise overview of Raz's theory and criticisms thereof, see Kenneth Ehrenberg, "Joseph Raz's Theory of Authority", *Philosophy Compass* 6 (2011): pp. 884-94 and "Critical Reception of Raz's Theory of Authority", *Philosophy Compass* 6 (2011): pp. 777-785.

these dependent reasons, I'll call them preexisting reasons.³⁴ Preexisting reasons are reasons the subject already had before a command. Like good advice, authority must be based on the preexisting reasons.³⁵

Advice and authority are distinct, however. Advice gives content-dependent reasons to do as advised while authority gives content-independent reasons to do as commanded. Advice must be based on preexisting reasons, but in order to accept it, the advice must actually, accurately reflect those preexisting reasons.³⁶ You have no reason to accept bad advice (except instrumentally, for example in order to appease your advisor). The fact that the commands of a genuine authority constitute content-independent reasons, on the other hand, means that sometimes commands will bind subjects even when they fail to accurately reflect the preexisting reasons. If authority were like advice and had to accurately reflect them, authority would again be rendered useless. This is another version of the idea that some mistaken commands must bind.

Although some mistaken commands bind, an ideal authority would never issue mistaken commands. If there were such an infallible authority, it's clear why it would be rational to obey her commands. If her commands always accurately reflect *your* preexisting reasons, and you know that you are fallible but she is not, then you can always conform better to your reasons by obeying the authority than by following your own judgment about the balance of reasons. Actual authorities aren't infallible though. When

³⁴ Because of this, Raz calls this condition on authority "the dependence thesis", MF, p. 47.

³⁵ It is important to note that a subject's preexisting reasons are not necessarily self-interested reasons. When a military commander issues an order that puts her soldiers in harm's way, it doesn't mean she's making her decisions on the wrong reasons. Instead, we understand that the soldiers have a reason to risk their lives in pursuit of winning the conflict.

³⁶ As a pragmatic issue, it is often difficult to tell the difference between good advice and advice issued by a good advisor. In order to evaluate whether advice is good in its content, you would need to already know what your reasons are and the advice would be redundant. But this pragmatic difficulty does not mean you should accept bad advice.

can it be rational to obey a fallible person who will sometimes issue mistaken directives?

The answer is quite intuitive: when obeying that person *tends* to help you better conform to your preexisting reasons. While your judgment might be better in individual instances, on the whole you will conform better in the long run if you obey than if you follow your own judgment. This idea is captured in Raz's main condition on genuine authority, the normal justification thesis:

[T]he normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly.³⁷

The main question: why does fulfilling the normal justification thesis mean that A's directives constitute preemptive and content-independent reasons for B? First, the reason is content-independent because it is about the *tendency* of the authority's directives. The important point is that an agent has authority when "the alleged subject is *likely* better to comply". It is not that following this particular command will *actually, necessarily* lead the subject to act better. Again, this would render authority redundant and it wouldn't explain how authority's commands give content-independent reasons. Instead, it's a tendency to better compliance, based on the reliability of the authority's judgment in that domain.

Due to this, Raz terms his account the *service conception* of authority: authority exists and is justified just because it is in service of the subject, helping her conform better to her preexisting reasons. A somewhat flippant summary of the idea behind the normal justification thesis is this: if agents aim to maximize conformity with their reasons, the only condition under which obeying an authority could be rational is when obedience would

³⁷ Raz, MF, p. 53.

increase conformity.

When an agent fulfills the normal justification thesis, the subjects' preexisting reasons are preempted. For now I'm going to gesture at an explanation of this fact. In chapter three I provide an in-depth examination of exclusionary reasons and the notion of preemption, and there I will provide a fuller account of the connection between preemption and the conditions on genuine authority. The idea is that my preexisting reasons are preempted because the authority bases her decisions on those reasons and her judgment about those reasons is likely better than mine. Like advice, genuine authority aims at serving subjects in a domain by issuing directives based on subjects' preexisting reasons in that domain. By virtue of representing a likely better balance of preexisting reasons, the command excludes those reasons. This sketch raises a variety of questions that I take up in in further chapters.

The normal justification thesis is importantly qualified by the independence condition: "that the matters regarding which [the normal justification thesis] is met are such that with respect to them it is better to conform to reason than to decide for oneself, unaided by authority."³⁸ Calling this the "independence" condition is a bit confusing because, if it is met, then subjects do *not* have a reason to act "independently", i.e. on their own judgment. Instead, I'm going to call it *the deference condition*, on the thought that if it is fulfilled, then the subject might have a reason to defer in this situation.

An example that fails the deference condition, so does not entail the authority of A over B although A fulfills the normal justification thesis with respect to B, is the process of responsabilization of children. In order to teach children how to make decisions on their

³⁸ See Raz, RSC, p. 1014; MF, p. 69.

own, parents often give their children discretion over decisions that the parents could make better. In this case the child making the decisions herself is more important than securing better outcomes by letting the parent make the reliably better decisions: the child has reasons not to defer. Similarly, in many areas of an individual's life it is more important that she choose for herself, exercising control over her own life and acting as the agent of her own decisions, than it is to secure better outcomes. Note that the deference condition merely captures contexts in which subjects do not have decisive reasons not to defer; whether a subject should actually defer also depends on whether the other conditions on genuine authority are met.

The main attraction of the service conception is its explanation of how obeying an authority can be rational. The normal justification thesis shows how the commands of an authority are connected to the preexisting reasons that already apply to the subject. It is only because of this connection that obedience is sometimes justified. Without the normal justification thesis, it would appear that obeying the authority is necessarily irrational, ignoring the reasons that apply to a subject and acting for the indirect reason of the command. But on Raz's account, the command itself must be based on the preexisting reasons that apply to the subject, so deference to the command is less mystifying. If rationality generally requires one to maximize conformity with reason, then deferring in those cases where deference serves conformity can be rational.

At this point we have clarified both the concept of authority and a plausible and widely accepted conception of authority. A has the power of authority over B in domain D iff A's directives constitute preemptive, content-independent D-reasons for B, with the result that some part of B's D-standing is changed. On the service conception, then, A's

directives constitute preemptive, content-independent D-reasons when following A's directives likely helps B better conform to her preexisting D-reasons. The goal of this extended discussion of authority is to help us clarify the debate between the philosophical anarchist and the governmentalist. Next in chapter two I apply our new understanding of authority to the problem of political authority.

2 – Evidential Modern Political Authority

From the arguments in chapter one, we now understand authority as the discretionary power to issue directives that constitute preemptive, content-independent reasons for others, resulting in a change in their standing in the authority's jurisdictional domain. When A has authority over B, A has the power to bind B at her discretion. In this chapter I apply this understanding of authority to the question of political authority. Political authority is differentiated from other kinds of authority in a variety of ways, the combination of which means political authority is particularly difficult to attain. That said, the broadened notion of authority allows us to see that there are different kinds of political authority. The standard debate between anarchists and governmentalsists has focused too narrowly on traditional authority, the right to command that creates duties. It is implausible that governments have traditional political authority, but plausible that they have another kind, what I call *evidential* political authority.

1. Political Authority

My aim in this dissertation is to examine the notion of political authority and attempt to adjudicate the debate between anarchists and governmentalsists in a new way, rather than to present an account of authority generally. The authority that governments claim has some necessary and distinguishing features that put conditions on genuine political authority beyond the conditions for authority in general.

Governments, as I defined them in chapter one, are the individuals filling institutional roles in a political institution, one prominent example of which is the nation-

state. For our purposes, there are three characteristics of political authority that distinguish it and call out for justification. The first two are intuitive and will be relevant later in the dissertation. The third requires more explanation and is one of the foci of this chapter.

First, political authority is diachronic. We can make sense of the notion of synchronic authority, authority in only one instance. When I promise to obey your directives about where to go to lunch, I make you an authority over only one domain and for only one act: where I go to lunch today. You are a synchronic authority over me. But this pushes the notion of authority to its limit; our normal understanding of authority is diachronic. This is why it makes sense to focus on the likelihood of following an authority's commands helping you conform better to your reasons over the long run. Governments claim a particularly diachronically robust form of authority because they claim authority over the entire course of subjects' lives. Laws function as *standing* orders.¹

Of course, we shouldn't close off the possibility that nobody can have authority that lasts that long. Perhaps governments' claim to authority over individuals' entire lives is simply mistaken. But there doesn't seem to be any principled reason why this should be so. Further, governments couldn't accomplish the aims they are required for, like securing justice, establishing the rule of law, and so forth, without diachronic influence of some sort. Governments also claim diachronic authority in the sense that they claim to bind subjects with commands that were issued before those subjects were born. In the United States, there are laws hundreds of years old that people are still punished for breaking. If genuine, this kind of far-reaching authority is quite astonishing. We need very good reasons to

¹ H.L.A. Hart, *The Concept of Law* (Oxford: Oxford UP, 1961): p. 23.

expect that one agent could have this sort of power over others. The robustly diachronic nature of political authority calls out for special justification.

Second, political authority is held over more than one individual. In chapter one I defined authority as a three-part relation between two agents and a domain. It is only by way of summary that we speak of one agent having authority over more than one person, or over more than one domain. Political authority is concerned with solving social issues, issues that arise in the context of multiple individuals interacting. In order to solve those issues, it must have authority over multiple individuals. It also necessarily requires authority over multiple domains. Here I'm simply noting that if authority over one person is difficult to attain, authority over many others should be that much more difficult. It is another feature of political authority in particular that requires extra justification.

Finally, and most importantly for both the justification of political authority and for our purposes in this chapter, political authority is *morally punishable* authority. Political authority is a form of moral and so practical authority: it changes subjects' moral reasons for action. To see that political authority is a species of moral authority, consider: if you are standing in your kitchen one morning, drinking coffee, and suddenly armed men burst into your home, handcuff you, and physically carry you into a small, locked room where you are held for the remainder of your life, you normally have a (justified) moral complaint. You have the right to resist being taken away and the right to escape when you are held. When it's not the government taking these actions we call them assault and kidnapping, recognize their impermissibility, and punish people for them. But when the government does this within the bounds of the law, they say you have no moral complaint. They claim you have no moral complaint because they issued a law that changed your moral standing, which you

then broke. Further, they claim that you have no moral right to resist their punishments, and will punish you further if you do. This is striking and clearly distinguishes the attempted exercise of authority from the raw exercise of power; not even the boldest of criminals would claim that you act *morally wrongly* if you resist them (stupidly, maybe, or imprudently, but not wrongly).

The government's authority plays a crucial role in explaining why citizens aren't wronged by punishment (if they aren't). The government only claims the permission to punish you if you disobey their genuine commands, e.g. if you break the law. Disobeying the government must be a morally punishable violation, then, and any good account of political authority must explain this. If disobeying some authority is a morally punishable violation, it is *punishable* authority.

Whether the government has punishable authority is an important question but it is also importantly distinct from the question of whether a particular person is wronged by a particular act of punishment. Just because I've done something punishable *in principle* doesn't mean that a particular act of punishment is justified. Punishment must be inflicted on someone who has committed a punishable violation in order to be all-things-considered permissible but other conditions must be met as well.

For example, punishment must be proportionate in order not to wrong the punishee. If I steal your phone, I have done something punishable. But if I am killed as punishment for this theft, I am wronged despite the fact that I was punished for a punishable violation precisely because death is a disproportionate punishment for minor theft. If some government has a legal system with wildly disproportionate penalties, then even if they had punishable authority, actual acts of punishment by that government would

be all-things-considered impermissible. While their citizens would be punishable in principle for breaking the law, all the actual punishments inflicted by the state would wrong the punishee because they violate the proportionality constraint. So when concerned with punishable authority, we are simply inquiring whether disobeying a command of this authority is a punishable violation in principle. The further questions that must be answered to determine whether the government has all-things-considered permission to punish its citizens for breaking the law fall under political legitimacy as I conceive it.²

Political legitimacy is about whether the government has the right to rule. I think genuine political authority is a necessary but not sufficient condition on political legitimacy. The right to rule requires authority but also much more, including things like a fair and proportionate institution of punishment. When distinguished in this way, it is clear that authority and legitimacy are distinct ideas. Justifying the power to bind others requires very different considerations than justifying the right to rule.³

2. Modern Political Authority

In addition to focusing only on political authority, I want to focus on only one species of political authority, which I call *modern* political authority. I am concerned with modern political authority because I want to answer the question whether anything like a modern government could have anything like the authority they claim to have. This is not due to an

² As with many concepts in such a long-standing debate, “political authority” and “political legitimacy” get used in a wide variety of oft-conflicting senses. I am simply stipulating that political authority is the question of whether the government has the power to change our moral standing by issuing commands. Political legitimacy, to my mind, includes much more, such as the permission to set up a coercive legal system. I am not interested in terminological debates, but have reasons to think this is a sensible and useful way to draw the distinction.

³ Cf. Bas van der Vossen, “On Legitimacy and Authority”, *Res Publica* 14 (2008): pp. 299-302.

inherent conservatism, assuming that governments *do* have the power they claim and then trying to give a justification. The space for critique and improvement of actual governments remains. Indeed, my conclusion will be that almost all actual governments do not have the sort of sweeping authority they claim.

Instead, my motivation is pragmatic: if there is any genuine political authority in the world today, it has to look something like the authority actual modern governments claim. My interest in authority flows from my interest in the wellbeing of individuals, and so focusing on the only kind of (political) authority that affects actual individuals makes sense. The focus on modern political authority is also shared with the most common form of philosophical anarchism, which makes an empirical claim about the sort of authority actual governments claim or could plausibly claim given the current state of the world.

Although my account will not justify the precise sort of political authority that modern governments claim, there may be genuine political authority that is significantly similar. It might be true, for example, that some of the laws of some governments bind most of the people in a territory most of the time. Even though this isn't the sort of perfectly general authority modern governments claim, it would be an important kind of authority to justify. Some problems can only be solved by large-scale coordinated efforts, and binding people such that they can be punished if they don't participate is necessary to coordinate. If those people couldn't be so bound, those problems would probably go unsolved.⁴ Any optimism on this front depends on the possibility of something like genuine modern political authority.

⁴ Philosophical anarchists often argue that private institutions can solve these problems, rendering the government unnecessary for the ends governmentals generally appeal to. For a recent argument to this

As with political authority, there are three features of modern political authority to highlight. Again, the first two are straightforward and will be relevant later in the dissertation. The third feature is most important and distinctive, and will be important throughout this chapter. The three relevant features that characterize modern political authority are scale, scope of jurisdiction, and lack of consent.

First, modern governments claim to have authority over millions and even billions of people. Above I said political authority is over more than one individual and that having authority over multiple people is harder to achieve. My point about modern political authorities here is twofold. Simply given the number of people in modern political units, the worry about having authority over multiple people is exacerbated. As importantly, these people are incredibly and increasingly diverse. This trend cannot be realistically reversed, even if it were desirable to reverse it (which it isn't). Modern political authority is addressed to pluralistic groups that disagree about many of the most fundamental facets of life, and a good account needs to acknowledge the fact of reasonable pluralism.⁵

We could imagine authorities that were political in some sense but didn't have this problem with scale and pluralism. If a village elects a leader and sets up some laws, the leader may have a kind of authority that can bind a village but couldn't bind millions. For example, if authority depends in part on trust, as I argue in chapter four, then the fact that a village leader could establish extensive interpersonal trust with each member of the village while a leader of a modern government cannot do the same with each member of her state leads to some differences in their authority. This is one way in which my account does not

effect, see Michael Huemer, *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey* (New York: Palgrave Macmillan, 2013).

⁵ John Rawls, *Political Liberalism* (New York: Columbia UP, 1993).

attempt to capture all kinds of political authority. I am concerned with the possibility of genuine political authority that could bind the sort of massive political groups we find in the modern world.

Second, modern governments claim jurisdiction over the most important domains in our lives, for example personal safety. Some jurisdictional claims of current governments will not be realized on my account, for example the claim to necessarily bind everyone within a territory. But any good account of modern political authority must explain how modern political agents can have authority over the domains of our lives we care most about. Modern governments also claim jurisdiction over particularly modern issues. For example, some modern governments' claim to authority over the domain of collective self-defense includes jurisdiction over nuclear weapons. No government in history has had authority over a capability that had the realistic possibility of ending all higher forms of life on Earth, as the US and Russia did during the Cold War. Authority over the standard capacity to wage war is one thing, but threatening all life on Earth is another. These are further respects in which political authority must be moral authority, but they also show that it must be moral authority of a particularly important type, as it addresses moral issues that are exceptionally weighty.

Third and most importantly, modern political authority is nonconsensual. Discussions of philosophical anarchism and governmentalism have historically focused in large part on the role of consent. This is understandable and laudable; one easy way to get authority over another is for them to consent to be your subject. Consent is the gold standard for permissibility: it distinguishes boxing from assault and borrowing from theft. Consent can make permissible what would otherwise be grievously wrong. If we are

worried about taxes as a form of institutionalized theft (or slavery!), a successful appeal to consent would be very helpful indeed.

I'll first explain how consent can ground authority in terms of my promise to go to lunch wherever you command. On a normal day where I have made no promises about my lunch venue, a wide variety of reasons count in favor of choosing a particular venue. I have a right to choose the closest venue, or the fastest venue, or the tastiest venue, or the venue where my friends generally eat, or the venue with the best combination of these traits. Distance, speed, tastiness, and companionship are all good reasons for me to decide between lunch venues. If you asked me why I was lunching where I was, I could cite any of these reasons as a good explanation. But when I make my promise, they no longer justify my action.

If you command me to go to Myrtle's and I decide to go to Joe's, I cannot cite Joe's tastier food as justification for failing to go to Myrtle's (even if it's an accurate description of my motivation for going to Joe's and so a good explanatory reason). This is precisely the sort of reason that promising is supposed to rule out, and if promising didn't rule out such minor reasons of taste it would be essentially useless. *Ceteris paribus*, on the day I've promised the only reason that justifies my choice of lunch venue is the fact that I promised to go wherever you chose and you chose Myrtle's.

Applying this framework to taxes: without consent, I have the right to use my income as I choose without your interference. You can't force me to save, or force me to buy good scotch, or anything else. You have a duty not to interfere with my decisions about my own resources. But if I consent to give you 20% of my income every year, then I am not at liberty to decide how to spend that 20%. By consenting I have waived my right to spend

it how I choose and now have a duty to give it to you, acting wrongly if I don't. The only way I can conform to the balance of my reasons with that 20% is to give it to you. If I promise to give you whatever percentage of my income you demand, you are a traditional authority over me with respect to my income because you can change my duty by issuing commands that change how much I owe you. This is essentially the government's claim about their authority with respect to taxes.

But as anarchists have shown, modern political authority is nonconsensual, so consent cannot explain how taxation (or any other coercive act) by modern governments is permissible. It is clear that universal actual consent is lacking in all modern governments, and plausibly always will be lacking. As Christopher H. Wellman has written, "only wishful thinking supports the belief that existing states have garnered the morally valid consent of all of their citizens."⁶ In light of this, governmentalsists have appealed to consent in other forms, from Locke's tacit consent to Rawls' hypothetical consent. Simmons argues, very convincingly I think, that none of these appeals work, largely because none of them can capture the power of actual consent to transform the impermissible into the permissible. Without consent, however, the idea that modern governments have traditional authority looks mysterious, as Simmons concludes.⁷

As with empirical philosophical anarchists like Simmons, I do not want to deny the possibility of genuine traditional political authority. We can imagine a small community forming with universal actual consent and thereby creating a government with traditional

⁶ Christopher H. Wellman, "Liberalism, Samaritanism, and Political Legitimacy", *Philosophy and Public Affairs* 25 (1996): 212.

⁷ There are of course other accounts of traditional authority that do not appeal to consent, for example those based in principles of fairness or a natural duty of justice. Anarchists also have a series of arguments against these positions, but I won't go into them here because my only point is that modern political authority cannot plausibly be explained by an appeal to consent. In chapter five I appeal to the natural duty of justice and defend that appeal against objections.

authority over them. But since I am worried about governments of the sort that we see in the world today, I will focus on the question of whether genuine nonconsensual political authority can be secured. In sum, then, an account of modern political authority must explain how an agent (individual or corporate) can have the power to issue morally punishable commands that bind large, diverse groups, without their consent, yet over some of the most important facets of their lives. Highlighting these conditions shows how odd modern political authority is and how it especially calls out for justification.

3. The Grounds of Punishment

My main focus for the rest of the chapter is the idea that modern political authority is punishable authority. For authority to be punishable, it must be morally permissible in principle to punish a subject for disobeying the authority. At a minimum this entails that the subject must lack a right against being punished for disobedience. If she did have such a right, others would have a correlative duty not to punish her. Punishing her would thus violate her right against punishment and wrong her. With the Hohfeldian framework from chapter one, we see that right and vulnerability are opposites, so lacking a right against punishment entails that the subject must be vulnerable to punishment. The commands of a punishable authority must make it the case that subjects gain a moral vulnerability to punishment for disobeying.⁸ Another way I'll put this is that modern political authority must be vulnerability-to-punishment-making authority.

⁸ Given this, it looks natural to claim that authority is the liberty to command. This is similar to the view of Robert Ladenson, "In Defense of a Hobbesian Conception of Law", in ed. Raz, *Authority*: 32-55. Ladenson claims authority is a justification-right (another term for liberty), so is similar in some respects to my account. But Ladenson's explanation of authority is much different, appealing to a Rawlsian hypothetical consent model, and is Hobbesian in that it basically says whatever government happens to be in power will have authority because everyone would recognize the necessity of the state from behind the veil of ignorance. I don't think either the substance of or the rationale behind this view are correct.

It is intuitive to think that traditional authority is exactly the sort of authority that one is morally vulnerable to punishment for disobeying. After all, traditional authority creates duties so disobeying a traditional authority is wrong. If we consider the most obvious cases of vulnerability to punishment, it looks like what we're punishing is wrongdoing, i.e. acting impermissibly and not doing one's duty. The Menendez brothers are in jail right now for brutally murdering their parents. They are clearly vulnerable to such punishment (punishing doesn't wrong them) and there appears to be a direct connection between their wrongdoing and their vulnerability. In murdering their parents they violated their parents' rights to life, wronged their parents, and failed to do their duty. In sum, they acted (egregiously) wrongly and it appears that this is what makes them morally vulnerable to punishment. But it turns out this very intuitive line of thought is mistaken.

Wrongdoing of the sort that disobeying the commands of a traditional authority entails is neither necessary nor sufficient to be morally vulnerable to punishment. To see this, consider Dr. Mary in the parallel cases of Benevolent Mary and Murderous Mary. Dr. Mary has a patient with a severe but not life-threatening dermatological condition.⁹ She has two treatment options: drug C, which will cure her patient, and drug K, which will kill her patient. Unfortunately someone has undetectably switched the labels on the vials containing the drugs, so C is labeled as K and K as C. Mary has a clear duty to give C and cure her patient, as well as a clear duty not to give K and needlessly kill her patient; the patient, in turn, has a right not to be unnecessarily killed by Mary as well as a (perhaps weaker) right to be cured. In the first version, Benevolent Mary gives the patient what she

⁹ Taken in spirit from Frank Jackson, "Decision-theoretic Consequentialism and the Nearest and Dearest Objection", *Ethics* 101 (1991): pp. 461-82. Also see Raz, FNR, pp. 120-6.

believes is C because she intends to cure her patient so selected the vial labeled C. But in fact this vial contains K so her patient dies.

Benevolent Mary acted wrongly: she needlessly killed an innocent and she failed to cure her patient when she had the means to do so. Killing the patient is inflicting an unjustified harm on an innocent and violates the patient's right to life, so he was wronged. But Benevolent Mary is clearly not vulnerable to punishment for giving K.¹⁰ (For now I'm simply going to appeal to intuition to establish this claim, as it seems significantly implausible to deny. Below I'll sketch a theory of punishment that explains the intuition.) Were we to punish her, for example by imprisoning her, she would have a justified moral complaint against us. In Hohfeldian terms, this is just what it means to say that she is not morally vulnerable to punishment, i.e. she has a claim-right against punishment, so we have a correlative duty not to punish her and wrong her if we do. The case of Benevolent Mary shows that not doing one's duty is insufficient for moral vulnerability to punishment.

In the second version, Murderous Mary hates her patient, so takes the vial labeled K and gives her patient a drug she believes will kill. Due to the label switch, however, her patient is cured and not killed. It is clear that Murderous Mary is morally vulnerable to punishment and is not wronged if we punish her. But we can't claim that she is vulnerable to punishment for not doing her duty and acting wrongly because she did her duty: her patient had no rights violated, was not wronged, and in fact has been cured of his disease. Instead Murderous Mary is morally vulnerable to punishment for attempted murder, an act

¹⁰ In case this seems too quick to you, note that we can easily stipulate all the conditions that make it clear she's not vulnerable to punishment: vials are generally accurately labeled in her experience, she has no reason to suspect she's not giving the cure, she has no reason to suspect the labels have been tampered with, and so on. As will become clear below, these conditions implicitly rely on the view that vulnerability to punishment is a function of responsiveness to one's beliefs and evidence.

worthy of punishment even when nobody's rights are violated and the agent happened to do what she should have done anyway. (This is comparable to a case where we punish someone for planning a murder, even though any of the individual components of the planning would be permissible without the intent to murder, like buying a weapon, watching a person in public, and so forth.) The case of Murderous Mary shows that not doing one's duty is unnecessary for vulnerability to punishment.

The difference between the cases that underwrites our opposing judgments of moral vulnerability to punishment is not the difference in outcome, or in rights violated, or whether Mary did her duty. Instead the relevant difference is what each Mary believed she was doing in light of her evidence. Benevolent Mary reasonably believed she was curing her patient when she gave K while Murderous Mary reasonably believed she was murdering her patient when she gave C. The fact that the outcome was the opposite of what they intended is irrelevant; Benevolent Mary and her patient were unlucky that the labels were switched while Murderous Mary and her patient were lucky. A similar way of putting this is that Benevolent Mary accidentally killed her patient while Murderous Mary accidentally cured her patient. Pointing out that the results were a matter of luck or accident also hints at an explanation why wrongdoing is not relevant for vulnerability to punishment: we only think it is appropriate to punish people for things that are within their control, and so things they are responsible for. We don't punish people for bad luck or accidents. Punishing Benevolent Mary for her bad luck is pointless and not punishing Murderous Mary for her good luck is foolish. But if wrongdoing is not the ground of vulnerability to punishment, what is?

4. Objective and Subjective

On the traditional account of authority, authority is duty-making. But traditional authority doesn't just create duties, it creates *objective* duties. It is common in normative ethics to distinguish between objective and subjective duties.¹¹ Objective duties are entailed by a decisive balance of objective reasons, which in turn are constituted by the relevant facts. Subjective duties are entailed by the balance of subjective reasons, which are constituted by something other than the facts, most commonly the agent's beliefs. (When I use 'duty', 'right', 'wrongdoing' and so on without qualifier I will always mean the objective type. If I mean anything else, I'll use a qualifier such as "subjective".) A perspicuous way of understanding the distinction is that objective reasons are not relativized to agents' epistemic standing while subjective reasons are so relativized. Your subjective reasons are what your objective reasons would be if your epistemic context was accurately reflective of the state of the world. If you are perfectly omniscient, your subjective reasons and objective reasons are always identical.

As an illustration I'll apply this distinction to the apple cases from chapter one. If I bought the apple, I have an objective claim-right to it and you have an objective duty not to take it from me. My right and your duty are objective because they hold regardless of whether we happen to believe they hold (they hold independent of our epistemic context, of which beliefs are an important part). If you happen to believe that the apple is unowned and take it from me, my objective right to the apple is violated and you objectively wrong me. You stole the apple from me regardless of whether you believed you were stealing it

¹¹ See, for example, Henry Sidgwick, *The Methods of Ethics* (London: MacMillan, 1874); Derek Parfit, *On What Matters* (Oxford: Oxford UP, 2011); Jeff McMahan, *Killing in War* (Oxford: Oxford UP, 2011). Note that in this context objectivity has nothing to do with moral realism.

and I have a justified moral complaint against you.

But your subjective duties are set by your epistemic context. If you reasonably believe that the apple is unowned, then you lack a subjective duty not to take it and don't do anything subjectively wrong if you take it. If your belief was true and the apple really was unowned, you would have an objective liberty to take the apple. Your subjective duty is determined by facts about your epistemic context rather than facts about my ownership of the apple: it is entailed by the balance of subjective reasons rather than the balance of objective reasons.

The commands of an objective authority, then, constitute objective reasons for subjects. The commands of a subjective authority constitute subjective reasons for subjects. It's important that just because something constitutes a decisive objective reason doesn't mean it's necessarily creates an objective duty, as I argued in chapter one. Recall that the fact that an apple is unowned is a decisive objective reason that entails that everyone has an objective liberty to the apple. This leaves open the possibility of objective authority that is not traditional authority, i.e. not objective duty-making authority.

The Mary cases established that objective duty-making reasons and objective vulnerability-to-punishment-making reasons are different. The objective duty-making reasons are the facts about the badness of death, the moral innocence of the patient, the needlessness of the patient's death, the doctor's options, and so forth. The balance of these reasons entails that Dr. Mary had an objective duty to give C and cure her patient as well as an objective duty not to give K and not needlessly kill her patient. Benevolent Mary failed to conform to these reasons because she accidentally gave K, so acted objectively impermissibly and wronged her patient. Murderous Mary conformed to the objective

reasons because she accidentally gave C, so didn't act objectively wrongly. But it was also clear that Benevolent Mary was not objectively morally vulnerable to punishment while Murderous Mary was objectively vulnerable to punishment. For this to be the case, the standard of objective vulnerability to punishment cannot be whether the agent conformed to her objective reasons. If it were, our judgments of objective wrongdoing and objective vulnerability to punishment would be identical and we would have to hold that Benevolent Mary is vulnerable to punishment for accidentally killing while Murderous Mary is not vulnerable to punishment for attempted murder. This would be bizarre.

Instead of objective moral wrongdoing, objective moral vulnerability to punishment is based on *subjective* moral wrongdoing: whether the agent conformed to her subjective reasons and did her subjective duty. Applying this thought to the Mary cases gives exactly the results we want. Mary has an objective duty to give C and not to give K due to the relevant facts. But because the labels were switched and Mary reasonably believes that C is K and K is C, Mary has a subjective duty to give the drug labeled as C and a subjective duty not to give the drug labeled as K. If her belief that the drug labeled as C is C were true, she would have an objective duty to give the drug labeled as C. Benevolent Mary does her subjective duty by giving the drug labeled as C, which unbeknownst to her and unfortunately is K. Murderous Mary, on the other hand, fails to do her subjective duty by giving the drug labeled as K, i.e. C. Our judgments of vulnerability to punishment track which agent did her subjective duty, not which agent did her objective duty. People are only objectively morally vulnerable to punishment for subjective moral wrongdoing.

Whether or not a person does their subjective duty grounds *blameworthiness* or

culpability.¹² Murderous Mary is blameworthy or culpable for her act while Benevolent Mary is not. Benevolent Mary committed a blameless wrongdoing. With this connection between subjective wrongdoing, culpability, and vulnerability to punishment, we can supplement our intuitive evaluation of the Mary cases with a theory of punishment. Following Larry Alexander and others, I think punishment aims to reduce unjust harms with a normative intervention, by giving agents' practical reasons not to perform acts that risk harm.¹³ Because punishment aims at changing people's decisions, punishment is only appropriate in response to acts that the agent chose for herself and was under control of, i.e. those she is culpable for. Permissible punishment requires *mens rea*, a guilty mind.

Punishing Benevolent Mary makes no sense because her decision-making process was good even though the results were bad. She was not culpable for her patient's death precisely because it was a result of factors outside of her control. Telling her that she has reasons not to kill her patient would not have changed the outcome because she didn't try to kill her patient. On the other hand, punishing Murderous Mary makes sense because she made a bad choice in light of her evidence, intending to unjustly harm her patient, even though the results were good. If you pointed a gun at Murderous Mary and threatened to shoot her if she kills her patient, her decision-making changes and she likely would no longer try to kill. This is the sort of effect we want punishment (and the threat thereof) to have.¹⁴ So it's not simply that our intuitions about the appropriateness of punishment don't align with objective wrongdoing in the Mary cases. If punishment aims to reduce

¹² This is a relatively common thought. See, e.g., Parfit, *On What Matters*, Vol. 1, pp.150-58 for similar points.

¹³ Larry Alexander, Kim Kessler Ferzan and Stephen J. Morse, *Crime and Culpability: A Theory of Criminal Law* (Cambridge: Cambridge UP, 2009). This doesn't entail an endorsement of deterrence theories of punishment; while the aim of punishment is reduction of harm, the justification needn't be.

¹⁴ Compare Warren Quinn, "The Right to Threaten and the Right to Punish", *Philosophy and Public Affairs* 14 (1985): pp. 327-373.

harm by giving agents reasons that effect their decision-making, it makes no sense to punish someone whose decision-making was flawless but whose results were bad because of circumstances outside her control, like Benevolent Mary.

One possible oddity in this discussion is that when we accuse someone of acting wrongly, we generally mean it with a certain degree of approbation. If someone acts objectively wrongly but subjectively permissibly, then she has committed a blameless or nonculpable wrongdoing. In this case approbation is inappropriate precisely because the agent is blameless. For this reason it is sometimes thought that we should only use “wrong” and its cognates in the subjective sense.¹⁵ I understand this impulse and am sympathetic to it but we shouldn’t jettison the notion of objective wrongdoing for several reasons.

First, objective duties are what subjective duties aim at: conforming to the objective reasons is the goal of action for the rational agent. It is what our beliefs and evidence are about, and when we discover that our beliefs are false or our evidence is misleading we modify our beliefs because we aim to conform with the objective reasons, not just what we believe our objective reasons to be. This is especially relevant when I am evaluating my actions *ex post facto* and realize what my objective duty was. If I did my subjective duty but failed to do my objective duty I may appropriately be regretful, and one way to express this regret is to be watchful for similar circumstances in the future so that I can do my objective duty then. (I shouldn’t feel guilty, though, because I did my subjective duty so am not culpable or blameworthy.) Even though I acted blamelessly there is room for improvement and as a conscientious and responsible person I should be looking for areas of moral

¹⁵ For a nuanced discussion of different ways ‘wrong’ is used see Parfit, *On What Matters*, Vol. 1, pp. 164-74.

improvement.

Second, all I have claimed so far is that objective wrongdoing is irrelevant for objective vulnerability to *punishment*. But punishment is not the only way we treat agents based on their actions. Objective wrongdoing is the main ground of objective vulnerability to *defensive harm*.¹⁶ This thought is common in the self-defense literature. Imagine I try to kill you because a madman has kidnapped my family and threatens to kill them unless I kill you. My attempt to kill you is objectively unjustified: you are morally innocent and have a right to life that I violate if I kill you. I wrong you and you have a justified moral complaint against me. But the madman has coerced me by threatening the death of my entire family and this coercion is a paradigmatic excuse, rendering me nonculpable for killing you. I commit a blameless wrongdoing. I shouldn't be punished for killing you because of my excuse (it's the madman who should be punished). But you do not wrong me if you shoot me in defense. I am objectively vulnerable to defensive harm because I am objectively wrong to kill you although I am not objectively vulnerable to punishment because my wrongdoing is excused.

The reason objective vulnerability to defensive harm is grounded in objective wrongdoing is that defensive harm is only justified when it stands a reasonable chance of succeeding in reducing an ongoing or imminent unjustified harm. For example, the police may shoot someone who is in the process of unjustifiably killing, but if he has just killed, is unarmed, clearly surrenders, and is no longer a threat to anyone, they may not shoot him in defense (there's nothing to defend). Now the question is whether he is culpable and so

¹⁶ See McMahan, *Killing in War*. McMahan uses 'liability' as I use 'vulnerability', but to my mind this is confusing because Hohfeld uses 'liability' as the term for one of the second-order disadvantages (which I haven't discussed here).

vulnerable to punishment for the killing. Similar thoughts are common in just war theory and the conditions under which (the very great defensive harm of) war is justified. Because defensive and punitive harm have different aims, though, they are justified by appealing to different reasons. Objective wrongdoing grounds objective vulnerability to defensive harm but it is subjective wrongdoing that grounds objective vulnerability to punishment.

The most important point to draw from this section is that there are differences between what makes an act wrong and what makes an act punishable. I've been putting this in terms of objective and subjective. I use these terms because they seem most natural to me and because they track the distinction between veridical practical reasons and non-veridical epistemic reasons. But you needn't agree with me that practical reasons are veridical or that what makes an act wrong is not doing one's objective duty. The insight that there is a difference in what makes an act wrong and what makes an act punishable, regardless of the nature of that difference, forces a wedge between traditional authority and what is required for modern political authority.¹⁷ In the next section I explore this gap and argue that it opens up the possibility of an overlooked kind of political authority.

5. Subjective Authority, Beliefs, and Evidence

Modern political authority is punishable authority, so it must be able to change what people are objectively morally vulnerable to punishment for doing. Someone is objectively morally vulnerable to punishment only if they are culpable. In turn, someone is culpable only if they don't do their subjective duty. For the government to issue commands that change what people may be punished for doing, then, it must change what they have a

¹⁷ For example, another possibility is that what makes an act wrong and what makes an act punishable are grounded in different types of subjective reasons rather than in objective and subjective reasons.

subjective duty to do. If the government issued a command that changed people's objective duties but didn't change their subjective duties, then if they disobeyed they would be acting wrongly but not culpably. This entails that they can't be permissibly punished. Further, we know that sometimes people can be punished for acting in ways that aren't objectively wrong, like Murderous Mary. So authority that changes objective duties isn't sufficient for modern political authority.

The question that confronts us, then, is what it takes to change subjective duties. As I have been using the idea, an agent has a subjective duty to act in some way if she would have an objective duty to act in that way if her epistemic context was accurate. So Benevolent Mary's subjective duty is to give the drug that is labeled as C but contains K because if her belief that the drug labeled as C is C were true, she would have an objective duty to give the drug labeled as C. Subjective duties are relativized to a particular agent's epistemic context. Changing subjective duties thus requires changing epistemic context. In particular, we are concerned with what changes in epistemic context entail a change in moral culpability.

There are two types of facts about an agent's epistemic context that are relevant for moral culpability: what an agent actually believes and what she should believe in light of her evidence. What an agent believes is obviously often relevant to culpability. If I aim a gun at you sincerely believing that it will fire despite the fact that it has been rendered inoperable, I am culpable for pulling the trigger even though I don't harm you and don't even have a chance of harming you. I believed I was going to unjustifiably kill you and that makes me culpable. One way to characterize this is to say that I failed to do my *rational*

duty.¹⁸ An agent has a rational duty to act in some way just in case she would have an objective duty to act in that way if her actual beliefs were true. This is an important mode of moral evaluation: it says something about my character when I act in a way that I sincerely believe is wrong.

To see that evidence is often relevant to culpability, consider another pair of cases with Dr. Mary. The facts about the patient and the doctor's options remain the same as the above, so the objective duties remain the same. This time the labels are accurate, but Mary's evidence about the effects of K and C is different. In the first case Data Mary has very strong (although misleading) evidence that K will cure and C will kill: many reputable studies have been published with this result, which she knows because she regularly attends seminars updating her pharmaceutical knowledge, and so forth. Just like Benevolent Mary, Data Mary sincerely believes that K will cure her patient, so has a rational duty to give K. Her evidence is very strong, so she also has an *evidential* duty to give K: if her evidence were accurate, she would have an objective duty to give K. When she gives K and her patient dies, she is not culpable just like Benevolent Mary. In both cases their evidence to give K was strong, there were no defeaters of the evidence, and they followed their evidence, as responsible agents should.

In the second case Anecdote Mary has anecdotal evidence that K will cure and C will kill (and no contraindicators). She sincerely believes this anecdote so her rational duty is to give K, just like Data Mary. Anecdote Mary is not under any particularly strong time constraint and she has a pharmaceutical guide in her pocket that clearly indicates K will kill and C will cure. Anecdote Mary doesn't consult her guide, so gives K and her patient dies.

¹⁸ Here I'm drawing on Parfit's distinction between fact-relative, evidence-relative, and belief-relative duties. See e.g. *On What Matters*, Vol. 1, pp. 34, 37 and 150-8.

Anecdote Mary should be evaluated identically to Data Mary with respect to the objective reasons as well as their beliefs, both of which were held constant between the cases. Data Mary and Anecdote Mary acted wrongly by unnecessarily killing their patients but both did their rational duty by taking the action they sincerely believed would save the patient. Intuitively, though, Anecdote Mary seems culpable in a way Data Mary is not.

The difference of course lies in their evidence. Anecdote Mary's evidence was quite weak, she knew the action she was taking was morally weighty, and she knew that she could easily improve her evidence if she simply checked the book in her pocket. She should have checked before she gave K and not doing so indicates a moral failing. Anecdote Mary is culpable and blameworthy for killing her patient in a way Data Mary is not: Anecdote Mary was *negligent*. Although Anecdote Mary believed she should give K, her evidence that supported that belief was weak and she had evidence that K will kill, so we can say she did her rational duty but didn't do her evidential duty. (Note that while Anecdote Mary is culpable to some degree due to her negligence, this doesn't mean she's as culpable as Murderous Mary. Attempted murder is generally more culpable than negligent homicide.) We morally evaluate agents based on their evidence in addition to the objective reasons and their beliefs.

Changing subjects' subjective duties requires changing subjects' epistemic context, so requires changing either beliefs or evidence. We are concerned with subjective moral duties, so are concerned with changing subjects' beliefs or evidence about their objective moral standing. Once we see this, we see that accounts of traditional political authority actually have a straightforward explanation for why subjects are objectively vulnerable to punishment for disobeying the governments' commands. Consider a subject deliberating

about what to do, searching for evidence of her objective duties. She is subject to a traditional authority, whose commands change her objective standing. As stressed in chapter one, authority must be exercised by a performance, paradigmatically an utterance. This performance is directed at the subject. So as the subject is deliberating about what to do, the authority enters and commands her to act in some way. The command constitutes a change in her objective duty, but it also constitutes evidence of her subjective duty. Under standard conditions, the utterance of a traditional authority is good evidence that one's objective duty has changed (under nonstandard conditions, this needn't be the case, as if the authority is being coerced, or if it is understood that she is joking around and doesn't intend to bind, and so forth).

Accounts of traditional political authority can exploit this close connection between the utterances of an authority and subjects' evidence of their objective standing. Raz, for example, includes a publicity condition on genuine political authority, requiring that governments make their commands public in order to bind subjects. It doesn't appear that Raz was motivated to include the publicity condition in order to establish a change in subjects' subjective standing, but it does have that effect. For some, the public command changes their belief. Generally this change occurs because they actually receive the command, e.g. they read the posting or hear the utterance, and because they believe the authority is genuine and binds them. If they didn't believe the authority was genuine, then the issuance of a new command would not change their beliefs about their objective duties because they don't believe that their objective duties can change by that agent's commands. If governments must change subjects' actual beliefs, many people wouldn't be culpable for disobeying the law: all those who were ignorant of the command and all those who didn't

believe the government actually had authority. If this were true, it would essentially vitiate governments' claim to any kind of general authority, and the government doesn't take ignorance of the law to be an excuse that removes culpability and vulnerability to punishment.

Luckily, though, subjects' epistemic context can change even if their actual beliefs don't change because subjects' evidence can change without their beliefs changing. Public commands that change objective duties plausibly change subjects' evidence, even those subjects that were ignorant of the command or those that don't believe the government has authority. When ignorant or mistakenly-anarchist subjects disobey the law, governments can claim to hold them responsible because the subjects' evidence entailed that they ought to do as the government commanded. The claim is that these subjects are like Anecdote Mary, having good evidence not to do as they did but ignoring that evidence and so acting negligently.

So the mere fact that citizens' subjective duties must change in order to be objectively vulnerable to punishment for disobeying governments' commands doesn't undermine accounts of traditional political authority in any serious way. Governments have objective duty-making authority and then they simply manipulate subjects' beliefs and evidence about their objective duties through normal means, like uttering a command or posting a law, and this changes what people are objectively vulnerable to punishment for doing. Simply pointing out that what makes an act wrong and what makes it punishable are different doesn't show that accounts of traditional political authority are mistaken. In the next chapter I'll argue against traditional authority. My point here is different. It is true that one way to change subjects' subjective duties is to change their objective duties in a

public way such that subjects' evidence of their objective duties also changes. The standard debate between anarchists and governmentalsists has been myopic because it ignores the possibility that subjects' evidence of their objective duties could change *without* changing their objective duties.

Changes in subjective duties are sufficient for changes in objective vulnerability to punishment. Consider Murderous Mary again. Mary has a subjective duty not to give the drug labeled as K: she both believes and has strong evidence that the drug labeled as K will unjustifiably kill the patient. When she gives the drug labeled as K, she is objectively morally vulnerable to punishment even though she ultimately cures her patient and conforms to her objective reasons. Now imagine that a hospital official comes in and tells Mary that someone has been switching labels, so the official switches the labels back such that they now accurately reflect the vials' contents. Mary sees this switch and believes the official, so her subjective duty has changed. She is now culpable and vulnerable to punishment if gives K, whereas before she was culpable and vulnerable to punishment if she gave C. The point is that her change in evidence was sufficient to change what she is objectively vulnerable to punishment for doing. Mary's objective duty never changed: it remains the case that C would cure her patient and K would unjustifiably kill her patient, and those facts set her objective duty. Changes in epistemic context can change what people are objectively vulnerable to punishment for doing even when objective duties don't change.

On traditional authority accounts, subjective duties are changed by changing objective duties and then manipulating subjects' epistemic context to co-vary with the changes in objective duties. The authority does not change subjective duties directly;

instead it changes subjective standing indirectly by changing objective duties in conjunction with a change in epistemic context. Now we see, though, that objective duties needn't change for a subject's subjective duties to change. This opens up the possibility that governments could change *only* subjects' evidence of their objective duties without changing the objective duties themselves.

My proposal, then, is this: modern governments lack traditional (objective duty-making) authority, as philosophical anarchists have claimed. But governments don't need objective duty-making authority in order to have punishable authority. Another, overlooked option is that governments have *subjective* duty-making authority: their commands constitute content-independent and preemptive subjective reasons for subjects, which entail a change in subjective standing. On this account, governments change subjective duties directly and without changing objective duties. When governments issue commands, the commands change what subjects are objectively morally vulnerable to punishment for doing even though they don't change what subjects are objectively morally wrong to do.

Before fleshing out this thought, it is important to note that subjective duty-making authority is not identical to objective vulnerability-to-punishment-making authority. I can only be permissibly punished if I am culpable, and I am culpable only if I fail to do my subjective duty. But these claims are 'only if', not 'if and only if'. There are conditions under which I fail to do my subjective duty but am not culpable, as when I have an excuse like coercion. The claim, instead, is that what I am *prima facie* objectively vulnerable to punishment for doing changes when my subjective duty changes. There are various defeaters between failing to do my subjective duty and being objectively vulnerable to

punishment, so the change is *prima facie*. My claim is not that any subjective duty-making authority also has objective vulnerability-to-punishment-making authority, but the reverse: if governments have objective vulnerability-to-punishment-making authority, they must have subjective duty-making authority as well.

I recognize that this proposal immediately raises a horde of questions. The idea that governments only have subjective authority is quite counterintuitive in some respects. I answer these questions, I hope, over the remainder of the dissertation as I elucidate the particulars of this proposal and answer some of the most common objections to it. The first question to address before these questions can be fully answered, or even fully asked, is, what kind of subjective duty-making authority could governments have?

There are two possibilities, given the two relevant features of subjects' epistemic context, beliefs and evidence. What an agent believes entails her rational duties; the power to change rational duties at an agent's discretion would thus be rational authority. An agent's evidence entails her evidential duties; the power to change evidential duties at an agent's discretion would thus be evidential authority. While these are both possibilities in some sense, as far as I can tell the idea of rational authority is incoherent.¹⁹

Imagine a "rational authority" issued a command to a subject: believe *p*! Rational authorities would necessarily change subjects' rational standing. But rational standing as defined is entailed by the agent's actual, current beliefs, i.e. what she would be required to do if her beliefs were true. Given this, if a command fails to change what the agent *actually* believes, her rational duties won't be affected. Whether the subject actually believes *p* as a

¹⁹ Compare Raz, RSC, p. 1034. As Raz argues, relying on a fine distinction between having authority over someone and being an authority, this doesn't mean that nobody is a rational authority in the sense that they know better what rationality requires. It's just that nobody has the power to change rational standing at their discretion.

result of a command is an open question. If she doesn't take the authority to have authority, she might not believe that p. If she is simply distracted and forgets whether the authority said to believe p or not p, she might not believe that p. If she cannot bring herself to give up her belief that not p, she might not believe that p. If a person's command doesn't constitute a decisive reason in some domain, though, she doesn't have authority in that domain.²⁰ Rational reasons are just actual beliefs and one person's commands cannot *constitute* a change another agent's beliefs (even if they sometimes do effect such a change). A normative power can't constitute changes in descriptive facts, like the subject's actual current beliefs, and so can't constitute changes in rational standing: there is no rational authority.²¹

Beliefs and evidence are the most relevant epistemic categories for morally evaluating agents relative to their epistemic standing. The agent's actual beliefs are relevant because there is a sense in which they are all she has to work from at the time of action and the agent's evidence is relevant because there are sometimes beliefs an agent should have but does not because of negligence. Given that there is no such thing as

²⁰ The only thing that necessarily changes beliefs would be a physical power, like if I brainwashed you into believing p. But brainwashing changes your normative standing in the same way my drowning in a shallow pond in your path changes your normative standing. Both change the descriptive facts first, entailing changes in the normative facts. Brainwashing changes your beliefs and this changes your rational standing, and my drowning in your path changes the objective reasons and so changes your duties. In neither case do I exercise any kind of authority over you because I don't change your normative standing directly, at my discretion and simply by expressing an intention to. My directives don't constitute decisive reasons within the domain. Instead what I am doing is exercising a physical power, at my discretion pressing a button on the machine that brainwashes you. The effects on normative standing are outside my control.

²¹ It is only true that the idea of rational authority is incoherent on this narrow sense of "rational", but it is a widely used and important sense because it captures minimal practical consistency. If there is no such thing as rational authority in this sense, it means that no agent can command another in a way that makes her necessarily open to rational criticism for failing to conform. Of course, we normally expect adult humans to believe propositions that have been made obvious, in part because we recognize that what we believe is not under our direct voluntary control. So often, or even most of the time, an authority's commands will change what others believe and so have an effect on her rational standing, it's just not a direct effect. It's not rational authority, but epistemic authority, the power to issue directives that constitute content-independent and preemptive reasons to believe, exercised over subjects that are well-functioning rational agents and so most of the time believe what they should.

rational authority, and absent any alternatives, if governments have subjective authority, it must be evidential authority.

6. Evidential Political Authority

An appeal to evidential authority as an explanation of how modern governments bind their citizens is only interesting insofar as it is distinct from both the philosophical anarchist and the traditional governmentalist. In the next chapter I deny traditional governmentalist by arguing that governments lack traditional authority. In order to deny philosophical anarchism, two claims must be established. First, that evidential authority is a coherent and possible form of political authority and, second, that modern governments actually have or can plausibly attain genuine evidential political authority. If the first can't be established, then we should be conceptual evidential anarchists, just as Wolff is a conceptual anarchist about traditional authority. If the second can't be established, then we should be empirical evidential anarchists, just as Simmons is about traditional authority. In the remainder of this chapter and the next, I aim to show that evidential political authority is a coherent idea. The final chapters go on to elucidate the conditions on such authority and then apply the theory to modern governments.

The main challenge to the conceptual coherence of evidential political authority arises once we see what evidential authority requires.²² When A has evidential authority over B in domain D, A's commands constitute decisive evidence of B's objective D-standing.

²² For views that also attempt to ground political authority in epistemic factors, see Donald H. Regan, "Authority and Value", *Southern California Law Review* 62 (1989): 995-1096 and Hurd, "Challenging Authority". This section clearly separates Hurd's view from mine; while both of us appeal to expertise and epistemic authority, Hurd thinks practical authority is a conceptually incoherent idea, whereas I only think the appeal to expertise works because in some contexts expertise grounds a kind of practical authority. Regan's appeal to indicators is more similar to my view, but he does not make the connection to culpability or punishment and so does not adequately situate indicator reasons.

The commands of an evidential political authority, then, constitute decisive evidence for subjects of their objective political standing. The idea is that there is some subset of moral issues that are the proper concern of the government. Call this subset the political domain. The political domain is a moral domain but does not encompass all of morality: for example, whether one should tell a white lie is a moral question but not a political question. There are some moral questions that the government has no business addressing.²³ The conditions on genuine evidential political authority include (but aren't exhausted by) those characteristics that entail that an agent's directives constitute decisive evidence about another agent's objective moral standing within political domains.

From the epistemology literature, we already have a good idea of what sort of agent can issue commands that constitute decisive evidence for others: epistemic authorities or experts.²⁴ The classic example is the scientist, an expert about some descriptive domain who can tell you what to believe in that domain, e.g. a physicist can tell you what you should believe about the physical makeup of objects. Notably, the testimony of experts changes your evidence in a radically different way than the testimony of non-experts. Non-expert testimony gives you some reason to believe, but you should balance that reason with your other evidence. It's like advice. Expert testimony, on the other hand, constitutes decisive evidence that preempts all your evidence that contradicts their judgment (within the domain of their expertise). You should believe what the expert says even when it flies in the face of all the evidence you have, even very strong evidence.

Consider that to my untrained judgment the table in front of me is a solid object

²³ This is captured by the deference condition.

²⁴ See e.g. John Hardwig, "Epistemic Dependence," *The Journal of Philosophy* 82 (1985): pp. 335-49, "The Role of Trust in Knowledge", *The Journal of Philosophy* 88 (1991): pp. 693-708, and Alvin I. Goldman, "Experts, Which Ones Should You Trust?," *Philosophy and Phenomenological Research* 63 (2001): pp. 85-110.

with no space between its parts. I have many years of experience with many different tables that give me evidence in support of my belief that this table is solid. Now, a colleague of mine whom I know to be an expert on physics tells me that the table is *mostly* space between particles and that my folk judgment about the table is radically mistaken. Her judgment on that issue is informed by a colossal amount of underlying theory and belief about the relation between experiments, evidence, and theory confirmation. Most of these facts cannot be related to me in a way that I can understand and use to balance the evidence on my own without extensive training and experience. Thus when the expert offers a judgment about an issue within her domain of expertise, I must defer and take her on her word. Whatever evidence I had to the contrary is overridden, and in this case it looks like I had a lot of reliable evidence from my experiences of the table, and many other tables and solid objects like it. My evidence is changed decisively by her expression of her judgment despite my considerable evidence to the contrary and the significant counter-intuitiveness of her judgment to me. Her judgment preempted my countervailing evidence and her expertise gave me a content-independent reason to accept her judgment.

This should come as no surprise precisely for the reason that experts are epistemic authorities. It would be quite odd if their directives didn't affect subjects in a way very similar to how I've been describing the effects of the directives of practical authorities.²⁵ So the essence of my proposal is that expertise is a necessary condition on genuine evidential political authority, indeed the main necessary condition. If governments' judgments are not expert, then they cannot have genuine evidential political authority. However, it is *extremely* important that I am *not* claiming that expertise is *sufficient* for genuine evidential

²⁵ Compare Raz, MF, p. 53, where he notes that it is likely that practical and epistemic authorities "share the same basic structure".

political authority. In later chapters I'll argue for the further conditions. For the remainder of this chapter I focus merely on the expertise condition, for the claim that expertise can ground *practical* authority under any conditions is extremely contentious.

Experts change evidence, but evidence can be misleading and experts can be mistaken. Reasons to believe and reasons to act are very different. Consider a case from the ethics literature, the question of whether I have a reason to drink from the glass in front of me filled with a clear liquid.²⁶ Because I think reasons for action have force regardless of epistemic context, I think you only have a reason to drink if that liquid will actually serve the end you are trying to achieve, regardless of whether you believe it will. If the liquid is gin and you have a desire for some gin, you have a reason to drink it. If the liquid is petrol, you don't have a reason to drink it because it isn't gin and won't fulfill your desire for gin. In chapter one I characterized this idea as the claim that practical reasons are veridical.

Evidence or reasons to believe, on the other hand, are not veridical. I can have completely decisive reason to believe something that is false. If a chemist tests the liquid in my glass and tells me it's gin, I have decisive reason to believe that it's gin. But if it's actually petrol and the expert was mistaken, the expert's judgment doesn't change whether I have a reason to drink because whether I have a reason to drink is entirely dependent on what's actually in the glass and is independent of what I believe to be in the glass. In terms from above, while I still have objective reason not to drink from the glass, my subjective reasons have changed because my evidence has decisively changed as a result of the expert's testimony.

²⁶ Bernard Williams, "Internal and External Reasons", in *Moral Luck* (Cambridge UP, 1980). I don't want to embroil myself in the motivational internalism/externalism debate, but if one is an objectivist in the sense I've defined it one is committed to motivational externalism about practical reasons. As I've stated, however, one needn't be an objectivist to endorse my account of political authority.

This introduces a problem for my account because if modern governments only have evidential authority, it looks like they are only changing my evidence and my reasons to believe but can't affect my reason to act. They are like the chemist: their judgment is decisive with respect to the evidence, but they cannot change the underlying reason to act that their judgment is evidence of. If they are mistaken and we follow their judgment, we will fail to conform to our objective practical reasons: we drink some petrol and fail to fulfill our desire for gin or in the case of moral reasons we act wrongly. Raz considered views like evidential authority under the broader heading of "recognitional authority" and raised a series of related concerns for them.²⁷ They all focus on this worry that giving subjects reasons to believe cannot ground practical authority because practical authority must give its subjects reasons to act. Political authority is a type of practical authority, so expertise cannot ground political authority.

Raz's argument rests on his claim that authority must make a "difference to what its subjects ought to do".²⁸ What Raz means by this is that authority must be duty-making. But given arguments from chapter one, I think a broader characterization is in order: in order to make a difference, authority must change subjects' objective standing. Traditional authority is objective duty-making authority and changes its subjects' standing by imposing new objective duties. Evidential political authority is subjective duty-making and so objective vulnerability-to-punishment-making: it changes subjective standing by changing subjective duties and changes objective standing by changing objective vulnerability. I admit that objective duty-making authority is in many ways the ideal form of practical authority. After all, evidential authority is trying to identify what a traditional authority

²⁷ Raz, MF, p. 28-31.

²⁸ Raz, MF, p. 48. Original emphasis.

simply creates by an expression of her will. But the fact that evidential political authority is not ideal does not mean it is not practical authority.

Putting the difference authority needs to make in terms of the need to give a “reason to act” is in some ways infelicitous because it seems to imply that the reason has to be in favor of performing some act. But practical reasons also determine which acts are merely permissible for us, without necessarily determining that one particular act must be performed instead of others. Let’s return to an example I used in chapter one, the police officer exercising her legal authority over you. She tells you to stop and later tells you that you may continue. The former command creates a duty for you, but the latter command doesn’t give you new legal duties and you don’t have any new legal reason to do or not do anything. Instead, you lose your legal reason to stop and you gain legal liberty. The officer’s second exercise of authority made a decisive difference to your objective legal standing. But this change was the introduction of objective legal liberties, not the creation of an objective duty. Only the broader notion of the sort of difference authority must make can capture this.

So authority must change standing but can change any component of standing without needing to be duty-making. It is also important that authority in a practical domain must change subjects’ *objective* standing in that domain. Changing subjective standing is just changing reasons to believe, so subjective authority in a practical domain still merely changes what reasons you have to believe certain propositions about that practical domain. This is the kind of authority that truly makes no difference to our practical reasons. All experts are subjective duty-making authorities. As I noted above, expertise is not the only necessary condition on evidential political authority. It is only

under certain conditions that not doing one's subjective duty makes one vulnerable to punishment, so it is only under certain conditions that experts can have evidential political authority. Under those conditions, their directives change both your subjective duties and your objective vulnerability to punishment. Because any change in your objective standing in a practical domain entails a change in your practical reasons, authority that changes your objective vulnerabilities is as much practical authority as authority that changes your objective duties.

Raz is also worried that if there can't be decisive practical reasons *in favor of* a particular act, as a duty creates, then governments cannot coordinate action. Coordinating action is a main task of modern governments and without such an ability they would look very different than we conceive of them. The paradigmatic case of "pure" coordination, where there are no moral reasons to favor either side so some method is required to get people acting in the same way, is which side of the road to drive on. It is permissible to drive on the right, as in America, or on the left, as in Britain, but it is impermissible for each person in large societies with ubiquitous driving to be at liberty to choose which side to drive on for themselves. This would lead to chaos and driving would be extremely risky. This is a pure coordination problem because while there is decisive moral reason for everyone to drive on the same side, there is no decisive moral reason for everyone to drive on a particular side. We need a convention people will follow but it doesn't matter if the convention is to drive on the right side or to drive on the left side.

If the government's political authority doesn't extend to actually creating objective duties, however, then how can they create a convention? If the government's commands can't make it the case that everyone has decisive reason to drive on one side, then it will be

permissible for people to drive on either side.²⁹ But it seems very clear to me that changing everyone's vulnerabilities such that everyone is objectively morally vulnerable to punishment (like a \$100 fine) for driving on the left while everyone was at objective liberty to drive on the right can effectively create a convention. If you have a permission to do A or B and no significant reason to choose one or the other, introducing a significant cost to A but not B means you will choose A almost all the time. This is sufficient to achieve the kind of conventions necessary for modern governments.

It may seem that this change in practical reasons is essentially the same as a highway robber pointing a gun at you and threatening you unless you act how he chooses. But impermissible threats and threats of punishment that one is morally vulnerable to have very different effects on the decision-making of an agent trying to act in perfect conformity with her practical reasons. If I am morally vulnerable to punishment, I have no right against it and so lose the right to resist. It is essentially a cost I have to accept if I choose the punishable action. But I do have the moral liberty to resist unjustified threats, creating different options for me and introducing a different balance of reasons. I can defend myself from unjustified threats, have a justified demand for compensation if such threats are carried out, and so forth. Making someone morally vulnerable to punishment gives them a practical reason importantly different from the reasons given by unjustified threats.

The idea of changing vulnerabilities to make conventions without changing duties means that it's not the case that you necessarily act wrongly in every case of driving on the

²⁹ It's actually not clear that traditional authority on the service conception fares any better here. This is a case of pure coordination, so *ex hypothesi* there are no reasons favoring either choice. But if there are no preexisting reasons for subject to choose a side, nobody will fulfill the normal justification thesis with respect to this choice because there can be no claim that following someone else's judgment could help a citizen conform better to her (nonexistent) preexisting reasons. It looks like Raz's argument depends on the *de facto* authority of the government to solve coordination problems.

left in America (assuming the American government has genuine authority in this domain). This seems true to me anyway, but whether driving on the left is objectively wrong is also both unnecessary and insufficient for objective vulnerability to punishment for driving on the left, as explained above. So objective moral vulnerability-to-punishment-making authority is practical authority and sufficient to give reasons for action that can create conventions on a large scale.

3 - How Authorities Bind

Evidential governmentalism claims that modern governments lack traditional political authority but have evidential political authority. In this chapter I give a novel argument against traditional modern political authority, and in the process more fully characterize evidential authority. The focus of the chapter is the decisive effect the commands of a genuine authority have on subjects' reasons, which I sketched but did not fully explain in chapter one. The sketch I gave followed Hart and Raz by explaining decisiveness in terms of preemptive, exclusionary reasons. While this sketch is essentially correct, I argue that exclusionary reasons have been misunderstood in some important respects. A better characterization of exclusionary reasons shows that they cannot ground traditional modern political authority, but can ground evidential modern political authority.

1. A Dilemma for Traditional Modern Political Authority

In this section I make a novel argument that modern governments lack traditional authority. In making this argument I proceed on the assumption that modern authority must be nonconsensual authority, as I argued in chapter two. Instead of arguing that particular attempts to ground traditional modern political authority fail, I argue that the idea of nonconsensual objective duty-making authority is incoherent. The problem is that accounts of traditional modern political authority cannot plausibly explain the effect of mistaken commands on subjects.

Consider Subject Mary, with the same patient and the same options as in the Mary cases from the last chapter. This means Subject Mary's objective duty is to give C and not to

give K, but in this case the vials are correctly labeled, so her evidential duty matches her objective duty. However, before she can give C an agent with genuine authority over Subject Mary in the domain of pharmaceuticals enters and commands Subject Mary to give K. If the authority has traditional authority, the command constitutes a preemptive and content-independent objective reason for Subject Mary that entails an objective duty to give K. Given the details of the case, this is a mistaken command: it commands Mary to do something other than what her preexisting objective reasons entail.

The problem is now apparent. Subject Mary has an objective duty to give C, entailed by the preexisting facts that C will cure her patient and K will kill her patient. But, according to accounts of traditional authority, Subject Mary *also* has an objective duty not to give C and to give K instead, entailed by the fact that a genuine traditional authority commanded her to give K. Subject Mary cannot have a duty both to give K and not to give K because of the “ought implies can” principle. Her duties are logical opposites such that performing one simply entails failing to perform the other: to give K entails not-not giving K, and not to give K entails not-giving K. The sense of possibility in the “ought implies can” principle I am appealing to in this case is very robust (and so very plausible): given the law of the excluded middle, it is logically impossible for Subject Mary both to give K and not give K, and so it cannot be the case that Subject Mary ought both to give K and not give K.¹

Before examining possible responses to this problem, first I want to note that it is not a problem on the evidential account because there is no conflict between evidential and

¹ One way to avoid this problem would be to deny “ought implies can”, as Raz does, FU, p. 1174. I will consider other responses because this seems implausible to me, in part because as I argued above the sense of possibility I am appealing to is very strong and so intuitively a plausible constraint on whether agents have a reason to act. But it is worth noting that this argument only applies to someone who wishes to defend the traditional authority of the government while also maintaining that “ought” implies “can”. Also note that maintaining that Subject Mary has both duties raises its own problems. The subject must fail in one of her duties and so necessarily acts wrongly, introducing a tragic dilemma.

objective duties. What you ought to do objectively and what you are culpable for as determined by your evidential duty are two different things, as the Benevolent and Murderous Mary cases illustrated. Raz is worried that this is still a problematic conflict in the sense that the subject must decide between her evidential duty and her objective duty but that there isn't a way for her to choose between them.² However, this is not a significant issue. I never denied that evidential duties are subservient to objective duties. Indeed, Mary's evidence is evidence about her objective reasons. The practical conflict cannot arise if Mary is aware that not giving K is her objective duty. If she knows that her objective duty is not to give K, she cannot have a conflicting evidential duty: her evidential duty is also not to give K. Her knowledge requires evidence and that evidence indicates that she should not give K. If she doesn't know this, then she acts wrongly but nonculpably if she gives K. Given Mary's epistemic context, she should follow her evidence and give K. But if we were advising her, we would tell her not to give K. These claims don't contradict each other. There is no practical conflict in the sense that we cannot guide her action or that she has no reason to choose one way of acting over another.

Returning to traditional authority, to avoid the dilemma the defender must show that one of the duties isn't present, either the preexisting duty not to give K or the command duty to give K. When comparing the duties, the most salient fact is that the authority's command was mistaken, so it is intuitive to claim that the authority's command failed to entail a duty for Subject Mary to give K. If one of the considerations has to give way, it should be the mistaken command and not the fact that giving K will needlessly kill Subject Mary's patient. The introduction of a command does nothing to change the medical

² Raz, FNR, p. 127.

and biological facts, and those are the facts that should determine what Subject Mary ought to do.³ This is all quite plausible, but we need to know why the authority's command failed to entail a duty in this case. After all, the case stipulated that the authority was genuine and imposing a new duty is precisely what traditional authority does.

The obvious thought is that the authority's command failed to bind Subject Mary just because it was mistaken. If the authority had accurately determined what Subject Mary ought to do and had commanded her to do that, Subject Mary would be bound. But given the mistake, she has no duty to obey. As initially compelling as this thought may be, it is not a good response. Recall Raz's statement that "there is no point in having authorities unless their determinations are binding even if mistaken".⁴ As emphasized in chapter one, if authorities only bind subjects when their commands are not mistaken, authority is empty and useless. The command of the authority adds nothing to the fact that Subject Mary should do what the objective reasons entail she ought to do, and if the command is mistaken she should just ignore it anyway. Authority isn't preemptive and doesn't seem to actually bind.⁵

Mistaken yet binding commands are made possible in part because of content-

³ This assumes objectivism about wrongness because the wrongmakers are not relativized to the agent's epistemic context. *The dilemma dissolves entirely if one is some an evidentialist about wrongness.* In that case the wrongmakers can be changed by a change in evidence: evidential authority would just be traditional authority.

⁴ Raz, MF, p. 47.

⁵ This is the same problem faced by the claim that while the command was genuine and gave Subject Mary a duty, it did not give her an all-things-considered duty. The commands of authorities, even genuine ones, only entail prima facie duties, and thus can sometimes be overridden. Any account of authority recognizes this fact, and appealing to it seems to solve our problem by claiming that the command did bind, prima facie, but was overridden by other concerns. Thus only the duty entailed by the preexisting facts actually (all-things-considered) binds Subject Mary and there is no problem with conflicting (all-things-considered) duties. However, in order to be a plausible response to the problem I have been pressing it would need to be the case that all mistaken commands are overridden in the manner just described. If the duties entailed by some mistaken commands are not overridden then they will conflict with the preexisting duty and the problem arises again. If all the mistakes are overridden, though, authority is rendered empty again.

independence: the reason to obey is the characteristic of the agent (her authority) rather than a characteristic of the content of the command (its accuracy). This contrasts authority with advice, where you should follow advice because it is good advice rather than because of who is advising you (although because of epistemic limitations, this distinction is often useless in practice).⁶ Whether the command is mistaken is an issue of content, but commands bind independently of content. This is why you could command me to lunch at either Myrtle's or Joe's: the content is irrelevant to whether it binds, and if you command me to go to Myrtle's I am bound to go to Myrtle's even if the balance of other reasons clearly favors Joe's, e.g. even if Joe's was closer, tastier, faster, cheaper, and where my friends eat. Content-independence is a necessary and defining feature the reasons the commands of a genuine authority give its subjects.⁷ Content-independence cannot be jettisoned and it entails that some mistaken commands must bind.

The defender of traditional political authority cannot plausibly claim that the authority's command did not entail an objective duty for Subject Mary. Perhaps, though, the authority's command really does change the situation such that the preexisting facts no longer entail an objective duty not to give K. Here I agree that the fact of a command from an authority adds a *new* reason for the subject to act in a particular way. But it is hard to see how this could eliminate the preexisting reasons. Sure, Subject Mary has a new reason to give K. But she still has decisive reason not to: the fact that giving K will needlessly kill the patient and violate his right to life. Both needless killing and violating rights are paradigmatic objective duty-makers. And she still has decisive reason to give C: the fact that giving C will cure her patient. The presence of the command does not, and could not,

⁶ Raz, MF, p. 52-4.

⁷ See Hart, "Commands and Authoritative Legal Reasons", and Raz, MF, especially pp. 35-7 and 46-8.

change these descriptive facts, and they still comprise reasons to act in certain ways.

This is not to say that nothing could change the situation such that Subject Mary doesn't have a duty to give C. Throw in a standard extreme consequentialist hypothetical (A terrorist threatens to blow up a nuclear device in a large city unless Mary gives K!!!) and the facts can change such that your duties change. But whatever change in the facts there is in Subject Mary's case, it is hard to see how the moral reason not to kill her patient could be overridden *merely by the fact of the command*. The reason given by the fact of command is not sufficiently weighty to override this kind of reason, where a life and the wellbeing of a person are at stake.

This is only a problem for the claim that modern governments, exercising nonconsensual control over morally weighty domains, have traditional authority. If the government has authority over a domain that is essentially trivial, as when I promise to obey your commands about my lunch venue, it is plausible that the weight of the mere fact of the command could override those reasons. But modern political authority is concerned with some extremely morally weighty domains that affect the wellbeing of hundreds of millions of people: establishing the rule of law, collective defense, the regulation of the market, the control of powerful destructive capacities, and so on.

Consent can explain how weighty moral reasons lose their force. Consider the following case: A has consensual sex with B and then is punished by his government for violating B's "family's honor". However, it is clear that A has an objective moral right against being punished for this reason and that the government wrongs A. Without consent, it is wildly implausible that the fact that defending honor is a legal justification could make it the case that the government doesn't morally wrong A when punishing him.

Defending honor is not a moral justification, particularly not when the so-called violation of honor is consensual.

But imagine that the government has secured universal actual consent from its citizens. A consented to waive his moral rights against the government punishing him as it sees fit; that is, he accepts that he is objectively morally vulnerable to punishment if he is legally vulnerable to punishment. Since the government considers violating honor a legally punishable offense, A accepts that having sex with B makes him morally vulnerable to punishment. It's as if A has said, "If I have sex with B, I have no moral complaint against the government punishing me", just like a boxer says, "If I enter the boxing ring, I have no moral complaint against my opponent if she punches me". Modern political authority is nonconsensual, however, so this line of reasoning cannot explain how it could be the case that A is morally vulnerable to punishment by any actual government, nor could it explain how the patient would not be wronged by Mary unnecessarily killing him.

The idea that a person's directive cannot override weighty reasons without consent is obviously true where the purported command comes from a non-authority. If Subject Mary is about to give C and is interrupted by a hapless bystander who orders her to give K instead, Mary should simply ignore the order. While requests by bystanders may indeed give some (very weak) reason to comply, this reason is clearly overridden by the reasons given by the life and wellbeing of the patient. But, the defender may say, the authority is not simply a bystander. There is something special about the authority that makes her authority genuine, and it is this characteristic that we should appeal to when explaining how reasons given by the fact of the command can defeat other, even significantly weighty, reasons. In particular, I have so far ignored the fact that authorities' commands constitute

preemptive reasons for subjects. If we are concerned only with the weight of reasons then this problem looks intractable. But preemptive reasons as I characterized them affect a subject's balance of reasons in a non-additive way, not merely as another reason to be weighed in the balance.

Raz's service conception tells just such a story: the directives of someone who fulfills the conditions on genuine authority constitute preemptive reasons for subjects because they help subjects better conform to their reasons. Because following the commands will help the subject better conform in general, Raz claims that the command of a genuine authority is preemptive in that it constitutes an exclusionary reason. So perhaps the reasons Subject Mary had to not give K are excluded. This would explain how the reasons not to give K do not entail a duty for Subject Mary and thus how Subject Mary can have a duty to give K despite the fact that the command to give K was mistaken.

In order to determine whether this move actually gets accounts of traditional modern political authority out of the dilemma, we need a fuller understanding of exclusionary reasons. While they are central to Raz's account, they have been challenged on a variety of fronts. In what follows, I address some of these challenges, sometimes defending Raz's account and sometimes admitting the force of the criticism and modifying Raz's account. In the end we will see that exclusion can't explain traditional modern political authority but fits well with an account of evidential modern political authority.

2. Excluded in What Sense?

Exclusion is a special mechanism by which some reasons are ruled out of the subject's deliberation, and it is only because of this special mechanism that we understand how

authorities affect subjects in the required manner. It turns out, however, that making the notion of exclusion precise is quite difficult. (Some have even claimed that the idea of exclusionary reasons is conceptually incoherent, a charge I address in Appendix A.) Raz's clearest comments come in response to Matthew Moore, who proposes three different interpretations of exclusionary reasons.⁸

The question that needs answering is what exactly is excluded and how. Moore's three interpretations of exclusion are the decision-procedure interpretation, the motivational interpretation, and the justificatory interpretation.⁹ On the decision-procedure interpretation, exclusionary reasons exclude certain decision-procedures from the subject's deliberations. On the motivational interpretation, exclusionary reasons exclude the agent from being motivated by certain reasons. And on the justificatory interpretation, exclusionary reasons exclude the justificatory force of certain reasons, making them no longer "count in favor of" some act. Moore claims that these interpretations form a trilemma for Raz because none works. In his response to Moore, Raz explicitly rejects the decision-procedure and justificatory interpretations in favor of the motivational interpretation.

To preview, I am going to argue that only the justificatory interpretation can answer the dilemma I proposed in the first section of this chapter. Raz and Moore both reject the justificatory interpretation, but that is because they are considering it in the context of traditional authority rather than evidential authority. The justificatory interpretation of exclusionary reasons given by the commands of an evidential authority is not only coherent

⁸ Raz, *FU*, pp. 1156ff.

⁹ Michael S. Moore, "Authority, Law, and Razian Reasons", *Southern California Law Review* 62 (1988): pp. 854ff.

but illuminating. That said, we shouldn't adjudicate among the interpretations merely on the basis of which can answer the dilemma. Instead, there are good, independent reasons to reject both the decision-making and motivational interpretations of exclusionary reasons. I'll explain why these two interpretations should be rejected before turning to the more controversial justificatory interpretation.

The first possibility is the decision-procedure interpretation, according to which excluded reasons are excluded as a matter of decision-making. When going through her practical deliberations, the subject cannot consider these reasons because they are excluded. But it is only this role in deliberation that the exclusion affects, and otherwise they act as reasons normally would. Hart seems to have embraced this interpretation,¹⁰ but Raz explicitly rejects it because he does not think authority affects subject's deliberation or what they can properly deliberate about.¹¹ When given a command, a subject merely has to act. If she wishes to deliberate about what she otherwise would have done, for example, she is free to. Moore adds a complaint to the decision-procedure interpretation, claiming that it would make exclusionary reasons merely first-order reasons to act on certain decision procedures, rather than about the reasons themselves. Taken together, these objections rule out the decision-making interpretation of exclusionary reasons.

The motivational interpretation of exclusionary reasons is Raz's preferred understanding, and the idea is simply that the subject is disallowed from acting on certain, excluded motivations. It is a claim about what the agent's actual psychological states

¹⁰ For example, in "Commands and Authoritative Legal Reasons", he says commands are intended to take the place of "any deliberation or reasoning" of the subject's own, p. 253.

¹¹ Raz, PRN, p. 48 and MF, p. 39.

should be at the time of action. This interpretation, however, has several difficulties. The first comes from Moore, who notes that people cannot simply choose which reasons motivate them. This is an analogue to doxastic involuntarism, which notes that people cannot simply choose what beliefs to have. Even if I offer you a million dollars to believe there is a giraffe in the room with you right now, you cannot just start to believe that. Similarly, one cannot simply flip a mental switch and find some reasons compelling and others not. If people cannot choose what reasons motivate them, however, then it cannot be claimed that people *ought* to be motivated by certain reasons over others (on pain of running into “ought implies can”). As noted above, Raz rejects “ought implies can” so can avoid this issue, but it is a significant worry for the overall plausibility of this interpretation.

Beyond the impossibility of simply choosing what reasons are motivating, our motives are also more complicated and more obscure than the motivational interpretation requires. On the motivational interpretation, the judgments of an authority make it the case that an agent is no longer allowed to act for those reasons. But most of the time our actions are taken for a variety of motives: I drove to work today because I was in a hurry, and because it was easier, and because I enjoy driving my car, and because I need to run errands after work, and because it is a habit of mine, and so on. All these reasons played a part in my actual psychology as I made the decision to drive. If an authority had ruled out one of those motives but not others, would I have been disobeying the command by acting as I did, multiply motivated? What if it only played a small part, but other motives would have been decisive in its absence? Not only do we often act for mixed, multiple motives, these motives are often obscure to us. This is clear both by introspection and in the social

psychology literature.¹² Just thinking about the mundane acts I took over the course of the day today, it is unclear to me what motives actually moved me, let alone what portion of my motivation for a particular act is ascribable to a particular motive.

For a command that excluded certain motivations to be effective, however, it seems that I need to be able to readily and accurately determine what I am motivated by and to what degree. Otherwise I have no idea whether I am obeying or not, and no idea what I can do to change so that I do obey. Thus commanding people to alter their motives seems difficult and very rarely effective, if not impossible.

The other problem with the motivational interpretation is that it just doesn't square with our understanding of authority, especially political authority. The odd part of this criticism is that it seems firmly grounded in Raz's own understanding of political authority. Hart claimed that authority changed how people ought to deliberate, not just how they ought to act. Raz distances himself from this claim, however:

Surely what counts, from the point of view of the person in authority, is not what the subject thinks but how he acts. I do all that the law requires of me if my actions comply with it. There is nothing wrong with my considering the merits of the law or of action in accord with it. Reflection on the merits of actions required by authorities is not automatically prohibited by any authoritative directive, though possibly it could be prohibited by a special directive to that effect.¹³

Here Raz is saying that because authorities only care how we act, they shouldn't care how we deliberate. But the same reasoning leads us to the conclusion that authorities shouldn't care about the content of our actual psychology when we act, i.e. it shouldn't care about our motivations.

¹² Richard E. Nisbett and Timothy D. Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes", *Psychological Review* 84 (1977): pp. 231-59.

¹³ Raz, MF, p. 39.

Consider the case of the spiteful private. The private is at boot camp under a particularly nasty drill sergeant. The drill sergeant wants to see this private fail, so orders her to do one hundred pushups. The private realizes that the sergeant intends to humiliate her by giving her an order she cannot obey. Upon this realization, she girds herself to do the pushups because she knows that if she succeeds, she will frustrate the sergeant's attempt to humiliate her. She does the pushups out of spite. If she didn't think she could spite the drill sergeant, she wouldn't do the pushups. When she finishes the pushups, has she obeyed? Clearly yes. If we embrace the motivational interpretation, though, we can't say this. On this interpretation, the sergeant's command constitutes an exclusionary reason that excludes reasons as motivations for the private, and one of the reasons that is motivationally excluded is spite. Her motivation was not the command itself, or any desire to obey, but precisely to spite the commander. Since spite was excluded by the command, though, she fails to conform to her reasons when she does the pushups and disobeys. The motivational interpretation gives counterintuitive results.

This is even clearer in the case of the law.¹⁴ I am planning to murder Jane, but ultimately decide not to. Consider three cases where I don't murder Jane but am motivated by different reasons: because there is a law against murder, because I think murder is morally wrong, and because it would just be too inconvenient. If motivation matters to the law, then I have committed a violation from the perspective of the authority if I don't murder Jane for any reason other than because there was a law against murder. But this seems preposterous especially when compared to the possibility that I don't murder only

¹⁴ See Larry Alexander, "Law and Exclusionary Reasons", *Philosophical Topics* 18 (1990): p. 15. Raz makes this point with respect to the law in one of the earlier writings: AL, p. 30.

because I think murder is wrong. If anything this seems like the ideal motivation not to murder, and we would feel oddly about someone whose only reason not to murder was the law. It clearly should not matter to the law that I don't murder because I think murder is wrong.

This shows, I submit, that it doesn't matter to an authority what subjects are motivated by and so shows the motivational interpretation of exclusion to be incorrect. Note that the claim that authority is unconcerned with motivation is not absolute. It is a live option that an authority issues a command with specifically motivational content. This does not show that the motivational interpretation is correct. The motivational interpretation claims that authorities are *conceptually* concerned with motivation; every command picks out motivations and there is no command from a genuine authority that is not concerned with motivation. The fact that the law doesn't concern itself with motivation thus provides a counterexample to the motivational interpretation. This conclusion is easiest to accept when the motivation seems noble, but it should extend to cases where the motivation is wrong-headed, like when I don't murder Jane because it would be too inconvenient. In that case the authority still has no complaint: it doesn't matter why I didn't murder Jane, the fact that I didn't murder Jane means I successfully obeyed the command not to murder. If we ask "Did I break the law?" the answer is clearly "No."

Further, it is unclear how the motivational claim fits into Raz's picture of authority and reasons. Authority is supposed to change what subjects' duties are as determined by practical reasons, i.e. relevant facts. Authority ostensibly accomplishes this by the fact that its commands entail exclusionary reasons. Exclusionary reasons are "second-order reason[s] to refrain from acting for some reason" and "Let us say that a person ϕ -s for the

reason that p if, and only if, he φ -s because he *believes* that p is a reason for him to φ .¹⁵ So the law against murder rules out not murdering because you *believe* that the fact that murdering is wrong is a reason for you not to murder. Surprisingly, this characterization makes it look like authority changes subjects' subjective reasons. An authority's command constitutes an exclusionary reason not to act on some second reason, but what it means to act on a reason is to act because of your beliefs about your objective reasons. The exclusionary reason thus excludes certain beliefs about objective reasons from being the reason the agent acts, but this is just to say that some subjective reasons are excluded.

Raz has a possible reply to these concerns. In earlier work he claims that what is excluded by the command of an authority is *all* the other reasons in the relevant domain.¹⁶ So when I'm considering murdering Jane, not only can I not act on my hatred and kill her, I can't act on my indifference and not kill her. The only motive available to me is to not kill her because the law said not to. And this is where the curiousness enters. But in later writing, Raz narrows the scope of exclusion.¹⁷ It is only motives that would lead one to act in a way other than how commanded that are excluded. So I still can't act on my hatred, because that would lead me to disobey the command, but I *can* act on my indifference or on my desire to do right because these motives would lead me to obey the command. Similarly, the spiteful private counts as obeying because acting on her spite leads her to do as commanded.

This move gets Raz out of the worry that his account is dissonant with our

¹⁵ Raz, PRN, p. 39, my emphasis. Raz's view has undergone many modifications over the years, so I don't take this understanding of exclusion from one of the earliest texts as necessarily definitive, nor do I take my concern based on it to be at all decisive.

¹⁶ Raz, MF, p. 59 and PRN, p. 192.

¹⁷ Raz, BAI, p. 144-5.

understanding of authority. But it does not escape the worry about the resistance of our motives to introspection and control and, further, it is quite odd.¹⁸ Notice that the motives really don't matter at all any more. Saying you can't act on any motive that would lead you to disobey is simply to say that you can't disobey. It has nothing to do with your motives whatsoever. The talk of motives simply operationalizes the claim that you must do as commanded. If you are motivated to murder someone out of hatred, that motivation is excluded for you. But if you are motivated the next day *not* to murder someone because of your hatred, you can act on it. The motivation is irrelevant; the act is all that matters. This is how we understand authority and putting it in terms of motivation is unhelpful.

A final possibility is that the counterexamples I raised are really not counterexamples at all. The thought is that if motivations really are so difficult to discern, then as a practical matter no authority can hold subjects to a motivational standard. Acting with a specific motivation is still required for obedience, but our practices of discovering when subjects obey and holding them accountable are different because of our lack of epistemic access to motivations. One way of thinking about this is to draw a distinction between the criterion of obedience and the decision-procedure for determining whether someone obeyed. The person who doesn't murder out of indifference really does break the law because she doesn't fulfill the criterion of obedience, there's just a pragmatic barrier to using the criterion itself to evaluate whether she obeyed. We use a different decision-procedure, namely just whether someone performed the act, because we need epistemic access to our evaluations. If the drill sergeant or the law had a way to discover people's motivations, things would be different. So my counterexamples simply highlight pragmatic

¹⁸ And as we see in Appendix A, this interpretation of the scope of exclusion runs into the problem of double-counting.

barriers and don't count against the motivational interpretation as such.

Ultimately, this is not an attractive possibility. First, it is still very counterintuitive to think that someone who didn't murder for reasons other than obedience to the law, for example the person who is simply motivated by their belief that murder is morally wrong, somehow disobeys the law. This is simply too restrictive. Second, it ignores the point that our motivations are resistant to introspection and direct intentional control and so the practice of holding us responsible for which motivation we act on is suspect. Third, it is at least worrying that the criterion for what counts as obedience and the methods for discovering obedience are so disconnected. In the end, the motivational interpretation may not be conclusively ruled out, but it is counterintuitive in a variety of ways. Whether we should embrace this interpretation ultimately depends on whether there are better alternatives. In the next sections I consider the justificatory interpretation and argue that it works, giving further reason to reject the motivational interpretation.

3. Justificatory Force

So much for the decision-procedure and motivational interpretations. The only remaining possibility on Moore's tripartite schema is the justificatory interpretation. I'm going to argue that the justificatory interpretation is correct. But this is difficult for a couple of reasons. First, both Moore and Raz reject the justificatory interpretation. Second, my main point will be that the justificatory interpretation looks exactly appropriate when we consider evidential authority. But as I have maintained throughout, I don't think the idea of traditional authority is conceptually incoherent, especially due to the possibility that genuine authority can be grounded in actual consent. If I argue that exclusion only makes

sense in the evidentiary context, then I'm stuck with the mistaken view that traditional political authority is a conceptually incoherent idea.

The justificatory interpretation claims that reasons are excluded in the sense that they lose justificatory force. As Moore characterizes it, "On this interpretation, promises and authoritative rules actually change what counts as a right-making characteristic of an action."¹⁹ The justificatory force in question is "right-making" force. Given distinctions I made in chapter two, this should be interpreted as objective duty-making force: the reasons no longer entail objective duties for subjects (I'll continue to use Moore's term 'right-making' as a convenient shorthand). As Moore goes on to say, "on this interpretation, some of the objective reasons of morality are excluded from doing their normal justificatory work whenever there is an exclusionary reason."²⁰ The talk of justificatory force could be translated into the different analogies we use to characterize reasons: for example, reasons lose their weight so no longer count in the balance of reasons when excluded. When reasons lose their justificatory force, they lose their character as reasons, for they no longer "count in favor of" anything. I'll continue to use the language of justificatory force because it clearly connects to the justificatory interpretation, but this locution may be confusing at times. To stave off this possible confusion, it is best to keep in mind that when I talk about justificatory force, all I'm referring to is reasons as we normally conceive of them, counting in favor of certain acts or beliefs and thereby justifying them.

At first glance, the justificatory interpretation is exactly what we need. After all, as Raz repeatedly emphasizes, authorities are supposed to change what the subject ought to

¹⁹ Moore, "Authority, Law, and Razian Reasons", p. 857.

²⁰ Ibid.

do. If they can change the justificatory force of reasons, then changing what subjects ought to do is simply entailed. Raz characterizes the exclusionary effect of an arbitrator's decision thusly: "The point is that reasons that could have been relied upon to *justify* action before his decision cannot be relied upon once the decision is given."²¹ If reasons can't be relied upon to justify action, they must have lost their justificatory force.

To me, one helpful way to adjudicate among the interpretations is with the distinction between explanatory and justificatory reasons. When we ask an agent why she performed some act, we could be asking two different questions about reasons. We could be asking for an explanation of her action, most commonly answered with a causal story. Or we could be asking for a justification of her action. These reasons often come apart. Imagine I bump into you on the street rather forcefully, causing you to drop your papers. When you ask me why I did that, I might explain that someone had just bumped into me and sent me stumbling into you. In that case I offer the fact that I was bumped as an explanatory reason, but I shouldn't pretend it's a justificatory reason. The fact that whether I hit you was out of my control excuses the fact that I caused you to drop your papers, rendering me nonculpable. But I wasn't *justified* in bumping into you; I caused you an inconvenience with no benefits to anyone. The fact that I was bumped first is an explanatory but not a justificatory reason, and *prima facie* I should be willing to compensate by helping you pick your papers up. On the other hand, maybe I bump into you because there is a piano falling from above and I wanted to knock you out of harm's way. In that case the explanation of my action is my desire to save you from harm and my justification for bumping you is the fact that if I didn't, the piano would have fallen on you.

²¹ Raz, MF, p 42. My emphasis.

Throughout the dissertation we have been concerned with justificatory reasons and not explanatory reasons. When a government issues a command, the interesting question is not whether the command explains my action but whether it justifies my action. We are concerned with *de jure*, not merely *de facto*, authority. More particularly still, the question is whether commands render disobedience unjustified in some sense. This is why we were concerned with whether commands entailed duties of obedience: duties, rights, and all the rest are relevant to whether actions are justified. A problem with the decision-procedure and motivational interpretations of exclusionary reasons is that decision-procedures and motivations are most commonly explanations of actions, not justifications. It looks like these interpretations are excluding certain reasons qua explanation instead of qua justification.

On occasion the decision-procedure I used or the motivation I acted on can be both my explanatory and justificatory reason. For example, when we are the ice cream shop and you ask why I chose chocolate rather than vanilla and I cite my preference for chocolate, my preference both explains and justifies my action. But as I argued above, commands are not conceptually concerned with motivations or decision-procedures. Instead, practical authority claims to change what actions of mine are justified and then lets subjects take those actions in whatever way they choose, by whatever decision-procedure and out of whatever motivation.

If the justificatory interpretation is so promising, why do both Raz and Moore reject it? Raz basically defers to Moore's criticisms of the justificatory interpretation since he endorses the motivational interpretation instead. All he says is that the justificatory interpretation would render some of his common examples of exclusion mistaken. I think

this is right but not an objection to the interpretation. This is especially true because I think we can still make sense of Raz's common examples once we concern ourselves with exclusion in the evidentiary context. I'll explain this below, after I consider Moore's objection.

Moore's objection to the justificatory interpretation is straightforwardly moral in nature and in fact can be understood as a version of the dilemma for accounts of traditional political authority I presented above. This is very interesting given that we are examining exclusionary reasons in such detail precisely to see if exclusion provides a way out of the dilemma. Recall that in the case of mistaken commands, it is difficult to see how authority "wins out". In morally weighty domains, the reasons that entail a preexisting duty for subjects matter a great deal. In Mary's case, what hinged on her decision was the life of her patient, so I objected that the mere force of the command couldn't defeat the weighty reason given by the patient's life. Commands can't make it the case that a person's life doesn't matter any more.

Moore makes a very similar complaint: "Morality never gives us reasons that exclude morally compelling considerations from counting to determine the rightness of keeping some promise or following some order. Morality never does this because nothing can be excluded in the balance of objective reasons of morality without leading to moral error."²² The "moral error" in Mary's case would be unnecessarily killing her patient; morality (personified) doesn't just ignore these sorts of goods. Nothing an authority could do could render someone's life less valuable. Appropriating something Moore says against the decision-procedure interpretation, "We do not have anything like this kind of

²² Moore, "Authority, Law, and Razian Reasons", pp. 872-3.

sovereignty over morality.”²³

In fact, this kind of intentional control over what constitutes objective moral reasons is so counterintuitive that we don't think even god could have such control. In metaethics the most common argument against divine command theory is a dilemma, standardly posed as a version of the piety dilemma Socrates raised in *Euthyphro*. On the relevant horn of the dilemma, we assume that god's commands simply constitute right and wrong. The gory point of this horn is that it allows morality to be whimsical and arbitrary: god commanded people not to murder but just as easily could have commanded people to murder. Because god's commands constitute right and wrong, murder would thus be required and not murdering would be wrong. But, goes the objection, morality *just isn't like this*. Murder *simply cannot* be required. Morality *simply cannot* be arbitrary. If authorities can issue commands that make unnecessary killing right, as in Subject Mary's case, morality is objectionably whimsical and arbitrary.

Given how close this criticism is to the objection I posed above for traditional authority, it should come as no surprise that I think Moore is correct here. People's lives don't cease to count as moral reasons simply because somebody issued a command, even a genuine command. But as I argued above, I can't take this line of thought too far because I want to claim that exclusionary reasons work in the evidential context. So to explain why Moore's objection to the justificatory interpretation is correct in some sense but incorrect in some sense as well, we need to narrow our focus.

Note that Moore specifically says that reasons like the life of another person can't be excluded in the justificatory sense “because nothing can be excluded in the balance of the

²³ Moore, “Authority, Law, and Razian Reasons”, p. 856.

objective reasons of morality”. Moore is worried about justificatory force as objective right-making or duty-making force. But there is justificatory force that isn’t objective right- or duty-making. Of course I am referring to justificatory force that is relevant to objective vulnerability to punishment. Further, there is justificatory force with respect to subjective reasons, and particularly with respect to evidence. Once we start making these distinctions we can see what has gone awry with Moore’s sweeping rejection of the justificatory interpretation.

4. The Justificatory Interpretation in Context

Ultimately arguing that the justificatory interpretation of exclusion works in different contexts isn’t a criticism of Moore, as he is concerned with whether the justificatory interpretation can make sense of Raz’s service conception. As far as that goes, Moore and I are in agreement: the idea that the commands of a nonconsensual authority exclude the justificatory force of objective duty-making reasons is morally objectionable. My point is that we can’t extend this criticism to other contexts. The idea that commands exclude the justificatory force of other reasons is not objectionable in either the case of consensual objective duty-making force or nonconsensual objective vulnerability-to-punishment-making force.

Consider first the case of consensual objective duty-making force. I have agreed that universal actual consent can establish genuine traditional authority, i.e. objective duty-making authority. If this is true, then consent must be able to make sense of the justificatory interpretation of exclusion in the objective duty-making context. Returning to the Subject Mary case, the problem with the justificatory interpretation of exclusion was

that the very weighty reasons not to kill her innocent patient can't just go away, especially not just because an authority came onto the scene. Morality isn't like that. She has preexisting reason to give C and not give K, based on the relevant biological and moral facts. When the authority commands her to give K, it purports to exclude those preexisting, countervailing reasons. But reasons of that sort can't lose their justificatory force just because someone says so.

That's almost right; those reasons can't lose their justificatory force just because someone *else* says so. But they can lose their justificatory force just because *the patient* says so, i.e. because of consent. The objective duty-making reasons are constituted about facts about the patient, the value of his life, his rights, and so forth. The justificatory force of these reasons can be excluded by the authority's command if the patient consented to let the authority have dominion over those aspects of his life in this instance.

We need to be careful about what excluding justificatory force precisely entails. The claim is not that the authority's command could make the patient's unnecessary death *good*. The unnecessary loss of the patient's life is bad regardless of whether he consents. There is a sense in which this badness retains justificatory force; even if someone has consented to let you treat them in certain ways, the fact that one of those ways is bad still gives you a reason to choose other options. But when worried about *authority's* effect on the justificatory reasons, we need to narrow our focus. In particular, the concern is about objective duty-making justificatory force. And it is clear that consent can affect objective duty-making justificatory force because of the simple fact that people can waive their own rights.

The thought is this: Subject Mary has a preexisting objective duty to give her patient

C. We're now imagining that *both* Mary and the patient have consented to the authority with respect to how Mary treats the patient. This is analogous to subjects promising to obey and defer to the government. Before the authority enters, Subject Mary has the same objective reasons as in all the other Mary cases. Upon the authority's command, though, the reasons Mary had not to give K are excluded qua their objective duty-making justificatory force. Without the consent, Mary is put into a bind by this command because the patient's right can't be overridden merely by the fact of the command. But if the patient has waived his rights against being treated as the authority prescribes, then the fact that the authority commands Mary to give K *can* exclude the justificatory force of the reasons Mary had not give K.

Notice that the authority is still required here. It's not that the patient waives his right against being killed generally. The patient waives his right not to have his fate determined by the authority's judgment, but up until that point Mary has a duty not to give K. When the authority commands Mary to give K, though, that duty is excluded and no longer has justificatory force. The exclusion is not objectionable because the patient waived his right and that explains why there are no longer any objective duty-making reasons for Mary not to give K. Given the patient's consent to the authority and Mary's obedience to the authority, she doesn't wrong her patient even though she needlessly kills him. The command of the authority explains why Mary's reasons not to give K are excluded, but it is the patient's consent that justifies the exclusion.

The idea that consent grounds the justificatory exclusion of certain reasons of course can't help accounts of traditional *modern* political authority because modern political authority is nonconsensual. Even if consent can cancel the right-making

justificatory force of certain morally weighty reasons, it is difficult to see how nonconsensual authority would not fall to Moore's complaint that those reasons are generally not subject to change at the discretion of an authority. Of course, I am not concerned with defending traditional modern political authority but evidential modern political authority. In order to maintain my claim that the notion of traditional political authority is not conceptually incoherent, I had to explain how justificatory exclusion makes sense in at least one case of traditional authority. Now that we see that notion of consensual authority depends on distinguishing between the command that excludes reasons and what justifies the exclusion of those reasons, we can better understand evidential modern political authority.

Epistemic reasons, evidence, justify in the sense that they count in favor of certain beliefs, rather than certain acts. If reasons are excluded in the sense that their justificatory force no longer counts, then when evidence is excluded it no longer counts in favor of certain beliefs. It is no longer evidence, just like a practical reason that no longer counts in favor of an act has lost its character as a reason. Recall that the objection to the justificatory interpretation of exclusionary reasons leveled by Moore was a straightforwardly moral one. Instead of the sort of conceptual objection leveled at the other two interpretations, Moore objected that it seems mysterious for right-making reasons to lose their force; it conflicts with our notion of morality as non-arbitrary. I agreed.

But it's not mysterious for objective vulnerability-to-punishment-making reasons to lose their force in this way precisely because objective vulnerability to punishment is a function of culpability and culpability is a function of evidence. Evidence, as I have been

stressing, is not veridical in the way right-making reasons are veridical. While a person's right to life can't change at another's discretion or any other person's, a person's evidence can change at another's discretion. It's not necessarily morally objectionable to accept someone else's testimony. (Although it is certainly more difficult in moral contexts.)

This is essentially the main argument I have been making throughout the dissertation. Morally, what is wrong for a person to do and what they are vulnerable to punishment for doing are different notions grounded in different reasons. How anyone could change what is wrong for another person to do is mysterious and does not accord with our understanding of morality. How anyone could change what another person is vulnerable to punishment for doing, however, is another matter entirely because of its connection to evidence. We have been discussing reasons and their justificatory force all along. Justificatory force is just another way of talking about the balance of reasons and whether reasons entail duties, liberties, and so on.

Excluding practical reasons is so difficult in part because practical reasons are veridical; the patient's rights are not subject to our discretionary control and he has them regardless of our evidence to that effect. But evidence is not veridical, and this crucial difference makes it plausible that evidence can be excluded in the justificatory sense even when practical reasons cannot. Here's another variation on the Mary case that demonstrates the claim. This time we are considering Android Mary. Android Mary has the same patient in the same predicament. Notice that Mary's duty to cure and not kill her patient are entailed by the right-making justificatory force of the reason that is constituted by the fact that the patient is a person with specific interests and rights.

Now, a technician enters the room and tells Mary that her patient is not in fact

human, but a very cleverly constructed android. Let's assume that androids are not moral agents or moral patients, but simply complex pieces of machinery no different in kind from toasters. The technician testifies to the patient's android nature, shows Mary a convincing data port behind a flap in the back of the patient's skull, shows Mary an android's blueprint, and so forth. The point is that Mary now has strong evidence that her patient is an android, not a person. This is evidence about her objective reasons, for her duties with respect to a human person and an android are very different. If the patient is an android and not a person, so has no rights, Mary does not have to treat him.

Before the authority entered, Mary had a right-making reason not to give K and good evidence that she ought not give K: her objective duty is not to give K and so is her evidential duty. The evidence given by the technician changes the evidential justificatory force of her evidence that she ought not give K; if the patient really is an android as the evidence supports, then it's not the case that Mary has an evidential duty to give C. Before her evidence changed, Mary would have been culpable if she didn't treat the patient. Now with strong evidence in favor of believing the patient is an android, she's not culpable if she doesn't treat him. Here we have a case where evidential justificatory force is decisively changed and so what Mary is culpable for changes, without changing the right-making justificatory force of any reasons. The right-making force is determined by the patient's true nature; if he's human after all and the technician was mistaken, Mary still has a duty to treat him even though she blamelessly believes she does not.

So it is at least *prima facie* plausible that evidence can be justifiably excluded in some cases where right-making reasons cannot be. Moore's moral objection to the justificatory interpretation of exclusion only holds in the objective right-making context. In

order to make this case more than prima facie plausible, though, we need a story like the consent story that appealed to something other than exclusion itself in order to understand how the exclusion is justified. That I find it intuitive in cases like Android Mary's doesn't yet show this. In the next section I look more closely at expertise and exclusion and show that exclusion is justified because of undercutting defeat.

5. Expertise and Exclusion

Consensual traditional authority is a coherent idea because consent explains how certain reasons can be justifiably excluded by the authority's commands: the subject waived her rights against certain kinds of treatment, so canceled the justificatory force of those rights and opened herself up to some of those reasons being excluded by the authority. Now we are looking for something to play the role of consent and canceling in the evidential context. This something is undercutting defeat.

Recall that a reason is undercut when its connection to what it counts in favor of is severed. In chapter one I gave the example where you find out that you are in fake-barn country, undercutting the connection between your visual perceptions of barn-like structures and the belief that barns surround you. The fact that you are in fake-barn country undercuts the evidentiary force of your visual perceptions because it provides a good error-theory for why your normally accurate visual capacities went awry in this context. Once you know that you are in fake-barn country, your visual experiences no longer count as evidence in favor of believing that real barns surround you: their evidential

justificatory force counting in favor of that proposition has been undercut.²⁴

Undercutting is based in considerations of reliability. The normally reliable connection between your visual capacities and believing what they appear to show is unreliable in fake-barn country, which is why the evidence from your visual capacities is undercut and no longer counts in favor of believing that you are surrounded by barns. Reliability is only an epistemic concern, though. If you happen to pass through the one part of fake-barn country that is filled with real barns, it is the case that you are surrounded by barns. But you still aren't justified in believing that you are surrounded by barns because you still have good reason to believe that your visual perceptions are unreliable in this context. The fact that undercutting is based in reliability and reliability is an epistemic concern makes it plausible that undercutting is functioning in the case of evidential authority. This is made even more plausible by two parallels to Raz: first, that Raz's favorite example of exclusion is actually a case of undercutting and second, the normal justification thesis is explicitly about reliability.

Raz's most common illustration of exclusion is drunkenness. The thought is that drunkenness affects your judgment, and knowing this, you should not act on your judgment. Drunkenness excludes the reasons you normally have to act on your judgment. Consider the case of the drunk text or drunk dial. During the hard, cold light of day you may have decided that you're never going to contact some ex-partner of yours ever again. But a few drinks in, you're seriously considering it; by the night's blurry end, you're punching in a number you swore to forget. Over the course of the night, the reasons to

²⁴ Recall from above that reasons lose their justificatory force very particularly. Your visual experiences in fake barn country still have evidential justificatory force counting in favor of believing that you are surrounded by movie props that look a lot like barns.

contact your ex start to look more convincing and the reasons not to contact look like they matter less. If you take your deliberations to give you a reason to contact your ex, you'll contradict your earlier, sober judgment. Should you?

Clearly not: you're drunk and drunkenness impairs your judgment. If you could sober up for a second, you would realize that your sober self made a good decision and that your understanding of the reasons is being confused by your drunkenness. Your drunkenness excludes acting on the fact that you currently judge contacting your ex would be a good idea. All well and good, except that this is clearly exclusion in the evidential, not right-making, context. The fact that you are drunk does not necessarily make it the case that reaching out to your ex is a bad idea. Whether that is a bad idea depends on a variety of factors like your current relationship with your ex, their feelings towards you, and so on. These facts aren't changed by the fact that you're drunk; it might be the case that your sober self was overly cautious so misjudged the reasons. Instead, the fact that you're drunk is good *evidence* that your current judgment is mistaken. The fact that you're drunk doesn't make your judgment mistaken, it just makes it more *likely* that it's mistaken. Drunkenness is a concern about reliability, an epistemic concern. Given that Raz uses drunkenness as the paradigmatic example of exclusion, then, it should come as no surprise that other instances of undercutting based in epistemic reliability can also ground exclusionary reasons.

Second, recall Raz's normal justification thesis:

[T]he normal way to establish that a person has authority over another person involves showing that the alleged subject is likely better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to

follow the reasons which apply to him directly.²⁵

Notice that the difference between the authority and the subject that gives rise to exclusion is the difference in how reliable their judgments are. As stressed in chapter one, in order to be content-independent, the difference must not be about whether their judgments are correct in a particular case. Instead, the difference is about whether their judgments are more likely to be correct in the long run, i.e. a difference in reliability. If I know someone's judgment is more likely to be correct than mine, that gives me good evidence that she is actually correct in particular cases where we disagree. But it doesn't make it the case that she's actually correct; differences in reliability constitute straightforwardly epistemic reasons. If this is right, then there must be some sense in which epistemic concerns of reliability ground exclusion.

At this point, it may seem intuitive to simply draw the connection between undercutting and exclusion like I did between canceling and exclusion. When a reason is canceled, it loses its justificatory force and then can be justifiably excluded by an authority's directive. Similarly, goes the thought, when a reason is undercut, it loses its justificatory force and then can be justifiably excluded by an authority's directive. The problem with this is that there is a crucial difference between canceling and undercutting. Canceling is binary: either a reason is canceled and completely loses its justificatory force, or it is not and doesn't lose any of its justificatory force. Undercutting, on the other hand, comes in degrees. A reason can be undercut to a lesser or greater degree and thereby lose its justificatory force to a lesser or greater degree.

Interestingly, Raz's example of drunkenness illustrates this point perfectly. Contrast

²⁵ Raz, MF, p. 53.

two cases where I am drunk and my inebriation gives me reason to doubt my judgment: in the first case I have had a couple beers and am buzzed while in the second I am extremely drunk, barely able to stand or form coherent sentences. In both cases I know my judgment is less reliable than it would be if I were sober, but clearly my level of drunkenness is relevant to how much less reliable I am, and in turn this is relevant to how much reason I have to ignore my drunken judgments. Consider: if I'm only slightly tipsy, then my slight impairment gives me some reason to doubt the reliability of my judgment, but my tipsiness may also clear away some inhibitions and allow me to consider options that I wouldn't otherwise, giving me some reason to think that my judgment is in some ways more reliable when tipsy. On the other hand, if I'm extremely drunk, slurring my words and unable to walk, my drunkenness gives me reason to believe that my judgment is completely unreliable. In that case I should not take the fact that I currently judge that I should contact my ex as any sort of evidence that I should contact my ex. My judgment when extremely drunk is so unreliable that it has been completely undercut.

Undercutting comes in degrees. Tipsiness is closer to the low end of the scale: the rational response is to lower my confidence in my judgment just a bit, and only temporarily. This doesn't mean I should never act on my judgment when tipsy, just that I am less confident and so need to be more careful, perhaps confirming decisions by deliberating again or by seeking advice. On the other end is complete undercutting. When I lose confidence entirely in a source I give no weight to its outputs: the reasons it gives to me have been completely undercut and no longer have justificatory force for me.

If we are appealing to undercutting as a parallel to canceling, then, it would appear that we have to consider cases of completely undercutting, rather than just any difference

in reliability whatsoever. Only complete undercutting removes justificatory force like canceling. At first glance, this is an area where Raz has gone wrong. Many have objected to the normal justification thesis on the grounds that it seems rather too permissive.²⁶ Note that the normal justification thesis contains no restriction on *how much* more reliable the authority's judgment must be in order to make her an authority. Any difference will do, even 0.01%, but this tiny difference in reliability does not intuitively seem to justify authority.²⁷

One implication of the fact that undercutting comes in degrees is that while concerns of reliability always give undercutting reason to some degree, they do not always give completely undercutting reasons. A small decrease in reliability gives rise to only slight undercutting; it would be irrational to entirely discount evidence that you suspected was only slightly mistaken. We can only explain exclusion by an appeal to cases of completely undercutting, where our confidence is totally undermined by unreliability. But this is what the normal justification thesis seems to deny. The normal justification thesis claims that A is an authority over B whenever following A's judgment would make B more likely to conform to the preexisting reasons than following her own judgment. Because authorities necessarily give exclusionary reasons, it must be that exclusionary reasons are given whenever there is any difference in reliability. The reliability claim is captured by "more likely to conform", and contains no limit on how much more likely is required. But

²⁶ For example, Moore, "Authority, Law, and Razian Reasons"; Alexander, "Law and Exclusionary Reasons"; Erman Mian, "The Curious Case of Exclusionary Reasons", *Canadian Journal of Law and Jurisprudence* 15 (2002): pp. 99-124; Kenneth Himma, "Just 'Cause You're Smarter than Me Doesn't Give You a Right to Tell Me What to Do: Legitimate Authority and the Normal Justification Thesis", *Oxford Journal of Legal Studies* 27 (2007): pp. 121-150; Stephen Darwall, "Authority and Reasons: Exclusionary and Second-Personal," *Ethics* 120 (2010): pp. 257-278.

²⁷ Although it may be that such tiny differences can be ignored if they can't be discovered or if one always has reasons to try to increase one's reliability that tiny amount (and so orders from someone who was only slightly more reliable would fall to the deference condition).

as we have just seen, this is wrong. Differences in reliability always give undercutting reasons, but they do not always give completely undercutting reasons. If exclusion can only be justified in cases of canceling or complete undercutting, the normal justification thesis is mistaken because it picks out many cases where reasons are only partially undercut.

If this is right, then we should modify the normal justification thesis to only capture cases where complete undercutting happens rather than any degree of undercutting. A completely undercut reason entirely lacks justificatory force, so we can easily explain why excluding completely undercut reasons is not morally objectionable. Better yet, it looks like an appeal to expertise is precisely what we need. The difference between experts and laypeople is a categorical difference in reliability. To see this, we need to know the nature of expertise more precisely.

6. Expertise

When we say a person is an expert, we generally mean something like she knows a lot about a particular domain. But expertise also seems to be more than just knowing many of the facts in a domain. The layperson may have quite a stock of true beliefs in the domain; for example, I can parrot a lot of interesting facts about physics and biochemistry. But for the layperson the justification for these beliefs is indirect, most often based in trust in a source like a professor or text. I believe Einstein's theory of relativity is true, but were I to go read the relevant physics papers and books I do not have the training or ability to competently assess them. My belief is not and, given my current limitations, could not be directly justified. If this is true, it would be odd to call me an expert about physics, even

though I know a lot, much more than the average layperson.

This is explained by John Hardwig's notion of an expert from the social epistemology literature. Hardwig claims that an expert is someone whose opinion is "less likely to be mistaken and likely to be less mistaken" than another because "sustained, prolonged, and systematic" inquiry is both "necessary and efficacious" in determining the truth of the matter.²⁸ This definition explains why just having more knowledge isn't sufficient for expertise. The crucial component is the inquiry because it establishes the difference between a knowledgeable layperson and the expert. Because of her inquiry, the expert has true beliefs plus much more. As Alvin Goldman puts it,

Expertise is not all a matter of possessing accurate information. It includes a capacity or disposition to deploy or exploit this fund of information to form beliefs in true answers to new questions that may be posed in the domain. This arises from some set of skills or techniques that constitute part of what it is to be an expert. An expert has the (cognitive) know-how, when presented with a new question in the domain, to go to the right sectors of his information-bank and perform appropriate operations on this information; or to deploy some external apparatus or data-banks to disclose relevant material. So expertise features a propensity element as well as an element of actual attainment.²⁹

The physicist, an expert, can go read Einstein's papers and form justified beliefs about the relation between the arguments therein and the conclusions, so can assess her belief in the theory of relativity. This allows her to change her beliefs in the face of new evidence, to extend the reasons for her belief to new beliefs, and so on. The layperson lacks this "propensity element" because she lacks understanding of the justificatory, second-order claims in the field. These second-order claims are things like the competing theories in the

²⁸ Hardwig, "Epistemic Dependence," p. 338, and n. 2.

²⁹ Goldman, "Experts, Which Ones Should You Trust?," pp. 91-2. Cf. Arthur Caplan, "Moral Experts and Moral Expertise: Does Either Exist?," in *If I Were a Rich Man Could I Buy a Pancreas?* (Bloomington: Indiana UP, 1992): pp. 18-39. On moral expertise's propensity element see p. 34.

field, the standards for good evidence, the appropriate methods for obtaining evidence, and so on. The expert's facility with these second-order claims, and her practical expertise at accessing and deploying her "information-bank", is a large and necessary part of what distinguishes her from the layperson.

Once we see that expertise is more than just a greater amount of true knowledge, we see that expertise grounds reliability of an especially robust kind. In particular, the expert has what I will call *trumping* reliability. To explore this idea it is useful understand this conflict in reasons as a kind of disagreement. Consider a case where the expert has told me I should believe p but I judge that I should believe not- p ; we disagree. The question is whether the disagreement gives me any reason to discount my own judgment, and if so, how much. This question is particularly difficult to answer when the disagreement is between peers. But when the disagreement is between a layperson and an expert, the answer is clear. The layperson should defer to the expert's judgment precisely because the expert's judgment is so much more reliable.

As Adam Elga characterizes it, when I treat someone as an expert I completely defer: "conditional on her having probability x in any [proposition within her domain of expertise], my probability in that proposition is also x ."³⁰ So if the physicist says she's completely confident that tables are mostly empty space, I should also be completely confident that tables are mostly empty space. I simply match her judgment of the relevant proposition because she has trumping reliability. At the same time, I must lower my probability in countervailing propositions, meaning that the reasons I had to believe those propositions are undercut. They are no longer good evidence that tables are solid, but have

³⁰ Adam Elga, "Reflection and Disagreement", *Nous* 41 (2007): p. 479.

been trumped by the expert's judgment. As Hardwig says, "The rational layman recognizes that his own judgment, uninformed by training and inquiry as it is, is *rationally inferior* to that of the expert (and the community of experts for whom the expert usually speaks) and consequently can always be *rationally overruled*."³¹

When an expert opinion gives a layperson reason to believe *p* and she also has reasons to believe not *p*, the mere fact of conflict with an expert opinion gives her a reason to lower her probability in those countervailing reasons. Her probability in the proposition supported by the expert is raised to the expert's probability, and her probability in other propositions is undercut so that they are at least lower than the expert-supported proposition. The balance of her evidence has decisively changed to support precisely what the expert believes.

Ultimately, though, we are not concerned with how evidence changes subjects' beliefs but with the conditions under which a change in evidence is sufficient to change a person's objective moral vulnerability to punishment. The appropriate epistemic response to changes in evidence depends only on the evidence, with the goal of having true beliefs. On the other hand, whether someone is objectively vulnerable to punishing for acting in some way depends on the combination of two factors: the strength of her evidence and the practical context. To see this, consider the chess case. In the chess case, you are playing a game of chess against some opponent with variable stakes. Let's assume that you're not a chess expert, although you've played some and are not a complete novice. You have an advisor on speed dial and can consult her as much as you'd like about how to play this game. You're considering move B(ishop) and call your advisor, who tells you that you

³¹ Hardwig, "Epistemic Dependence", p. 342.

should make move R(ook). If you make move B rather than move R, may you be punished? This judgment depends on two variables: the advisor and the stakes.

In half the cases, your advisor is a chess peer, someone you have played before and rightly consider to be of equal chess ability. In the other half, your advisor is a Grand Master. The Grand Master is a chess expert with respect to you, so her judgments constitute decisive evidence for you about how to win chess matches. First consider Fun Chess, where the only stakes are the satisfaction of winning. Whoever wins the chess game gets the satisfaction of winning and then everyone goes home. In the Fun Chess example, your advisor is irrelevant to your vulnerability to punishment. If you B, you may be irrational for ignoring your advice. But nothing much hinges on your decision, so you can't be punished if you B. In this case the stakes are so low that punishment isn't an appropriate option at all.

Second, consider Millionaire Chess: the winner of this game gets a million dollars. The stakes are much higher in this game, which might make it seem like your advisor becomes relevant again. But this isn't the case because although the stakes are high, they are not morally high; that is, whether you can be punished generally depends on how you treat others, and while gaining a million dollars would be good for you, it is not necessarily relevant to how you treat others.

To see this, contrast Millionaire Chess with Hero Chess. In Hero Chess the stakes are much different: if you don't win, the world is destroyed. You are considering B but your advisor tells you to R. If the world depends on your chess moves, may you be punished if you B? In the Grand Master Hero Chess case, it seems clear that the answer is yes. Imagine that the world was watching your game of Hero Chess. When the Grand Master tells you to

R and but you decide to B instead, do all the people whose lives depend on your moves have a justified moral complaint against you? Would they be justifiably upset at you for ignoring the Grand Master? Clearly yes. You are culpable for picking B and because the stakes are so high, harsh measures like punishment can be appropriate.

Contrast this with Peer Hero Chess: you have called your advisor, who tells you that you should R. Can you be punished if you B? It seems to me not. The testimony of a peer provides you with some evidence that R is the better option. But it's not decisive evidence; as I mentioned, how you should respond to peer disagreement is an extremely difficult question in epistemology. Luckily we don't have to solve this issue for our purposes here. The point is that however you respond, whether you continue to prefer your own view or whether you adjust your probabilities so you think it is equally likely that you should B or R, you can't be punished if you choose B. While the peer's advice gave you some reasons to believe that you should R, those reasons aren't decisive and so can't ground a change in vulnerability to punishment. In Peer Hero Chess it seems to me that you can either R or B and aren't vulnerable to punishment for either choice precisely because your evidence is so evenly balanced.

Whether a change in evidence changes what you are morally vulnerable to punishment for doing depends on whether your change in evidence is decisive and whether it is evidence pertaining to some morally weighty decision. All this comes down to, though, is whether you had an evidential moral duty: did you have decisive evidence that you morally ought to act in some way? In the Fun Chess and Millionaire Chess cases, you had decisive evidence that you should act some way when you called the Grand Master, but you were under no moral requirement. Not doing as the Grand Master says was

imprudent, maybe, but not morally wrong. In the Hero Case, it's clear that you have a moral duty to win the game. In the Grand Master case you also have decisive evidence that you ought to R if you want to win the game. Given the importance of the moral duty, if you B, you don't do your evidential moral duty and so are culpable and objectively vulnerable to punishment. In the Peer case, though, while it's still clear that you're under a moral duty, your evidence about R and B is inconclusive. You have no evidential duty and aren't culpable for either choice.

In Grand Master Hero Chess, you think that your best move is B based on some evidence. The Grand Master then tells you that you should R. Her expert testimony constitutes preemptive, content-independent evidence that undercuts countervailing evidence in the domain of her expertise, i.e. how to win at chess. Because it has this decisive effect, you have an evidential duty to R. Given the importance of the stakes, you are objectively vulnerable to punishment if you B. The evidence you had to B was undercut, so no longer counts in favor of believing that you should B, so no longer entails that you are not objectively vulnerable to punishment if you B. The undercutting was sufficient to exclude your evidence given the practical context.

Here's what this chapter has shown. The unique and characteristic effect of authorities' commands is their preemptive effect on subjects' balance of reasons and the ability to unobjectionably exclude other reasons' justificatory force. In some contexts this looks problematic, as in the objective duty-making context. Objective duty-making reasons are things like people's rights, which intuitively cannot be under another person's discretionary control without the right-holder's consent. In other contexts, though, exclusion is unobjectionable. This is true of the context we are concerned with, namely

objective vulnerability-to-punishment-making reasons. Because objective vulnerability to punishment is grounded in culpability and responsiveness to evidence, sometimes one person's objective vulnerability-to-punishment-making reasons can be decisively undercut. Due to its trumping reliability, expert testimony has this effect on laypersons' evidence. In some practical contexts, what one agent is vulnerable to punishment for doing can be changed at the discretion of another agent.

Thus we have come to *the expertise condition* on genuine evidential authority. Changes in evidence that are not decisive cannot give subjects an evidential duty and testimony that is not expert does not decisively change evidence. Raz's normal justification thesis is essentially correct, but instead of the subject just being "more likely" to conform to her reasons, the expertise condition holds that the subject must be much more likely to conform because she is a novice relative to the authority. The expertise condition is the foundation of my account, and I also adopt the deference condition from Raz. In the next chapter I more closely examine the sort of expertise and practical context required for genuine authority, ultimately adding three more conditions.

4 - From Expertise to Authority

To this point I have argued that expertise can sometimes ground the sort of moral authority that changes subjects' vulnerability to punishment. I now want to apply this picture of expertise and authority to the modern political context. Specifically, we need to know what kind of expertise is required for modern political authority and why it might be plausible that modern governments could attain such expertise. It is only if we have these details that we can go from the claim that evidential modern political authority is a conceptually coherent position to limited governmentalism, the claim that some actual governments have genuine evidential modern political authority. Otherwise we should just be empirical philosophical anarchists about evidential modern political authority, just as Simmons and others are empirical philosophical anarchists about traditional modern political authority.

In this chapter I argue for three further conditions on genuine evidential modern political authority: the acceptance condition, the trustworthiness condition, and the precedence condition. By the end of the chapter, we will have a complete picture of the conditions on genuine evidential authority: 1) expertise, 2) deference, 3) acceptance, 4) trustworthiness, and 5) precedence. Jointly these conditions are sufficient for genuine authority, although they are not necessary at least because of the possibility of genuine authority grounded in consent. After arguing for the latter three conditions, I go on to address why it is plausible that modern governments could meet the two conditions that are features of the authority, the expertise and trustworthiness conditions, which modern governments look particularly ill-prepared to meet. In the next chapter I argue that some

actual governments meet these conditions so have genuine evidential modern political authority.

1. The Difference Authority Makes

The characterization I gave of expertise in the last chapter, appealing to Hardwig and Goldman, is clearly a characterization of epistemic authority. Hardwig said an expert is someone whose opinion is “less likely to be mistaken and likely to be less mistaken” than another because “sustained, prolonged, and systematic” inquiry is both “necessary and efficacious” in determining the truth of the matter.¹ This allows us to see why the expert’s judgment is so reliable that it trumps the judgments of the layperson within the domain of her expertise. The sustained inquiry gives the expert a bank of information as well as a set of skills and techniques that allow her to make reliable judgments across a range of questions in her domain, what Goldman called the “propensity element” of expertise. But if our concern is practical authority, why am I focused on epistemic expertise? Shouldn’t I be focusing on practical expertise?

Practical authority is about telling others how to act. It is therefore intuitive to think that practical authority is grounded in practical expertise. If a good starting point for epistemic expertise is the person who knows many true propositions in some domain, a good starting point for practical expertise is the person who can very reliably successfully act in some domain. The expert archer, for example, can successfully hit her target very reliably. The practical expert would also have a propensity element, being able to successfully act in new and unforeseen circumstances.

¹ John Hardwig, “Epistemic Dependence”, p. 338.

The fact that someone has practical expertise is often relevant for decision-making, but not in the same way as epistemic expertise or authority. If the world is going to end unless someone hits a small target with an arrow then we should choose the archer. But authority is not about performing an act oneself, but about commanding others to act. And the fact that the archer can hit the target better than I does not entail that she can tell me how to shoot. Practical expertise is something like a finely honed skill, but practical expertise in a domain doesn't entail epistemic expertise in that domain. The archer may not be able to explain to others how she does what she does, but this doesn't undermine her practical expertise. Think of the autistic mathematical savant. She's a practical expert just because she can perform so well, regardless of whether she can articulate what she's doing and thereby tell others how to perform well.

Even if I know how to do something, I may not be able to explain how I do it to others. And even if I can explain how I do something, I may not be able to tell others how to do it successfully for themselves. As I stressed in the first chapter, authority is performative so requires more than good judgment, it requires the ability to transmit that judgment to others. If the archer can hit the target every time but can't successfully get others to hit it for themselves, she is a practical expert but can't be a practical authority.

Of course, this is not to say that practical expertise and practical authority are unrelated. Any actual practical expert is likely a practical authority about some aspects of her domain of expertise because any actual practical expert is likely a person who can successfully communicate about relatively uncomplicated matters with others, and some aspects of her domain can be conveyed in that way. Most practical authorities will also be practical experts. The person who can effectively direct others to cook is probably a good

cook herself. However this certainly needn't be the case. Imagine the eighty year-old hitting coach. She can teach others how to hit at the highest levels but is physically incapable of hitting herself. She is still a practical authority and epistemic expert about hitting even though she isn't a practical expert. You appeal to her if you want to learn how to hit a ball but not if you want a ball hit. Practical expertise and practical authority are related but conceptually distinct, and practical expertise is not necessary for genuine practical authority. (Practical expertise is sometimes a good *indicator* of practical authority.)

Epistemic expertise is required for practical authority, not practical expertise. So far, though, I have been characterizing evidential authority as if its effects were indistinguishable from expert advice. That is, expert advice also changes subjects' evidence decisively; that's just what it is to be an epistemic expert. What's different about evidential authority? First, there are many ways that an expert, even an expert in a relevant domain, could not meet the conditions on authority. Expertise is not sufficient for authority, as I have emphasized. This is clearest when we focus on authority's illocutionary aspect.

Usually when we engage with epistemic experts, we are asking their opinion about some proposition in their domain of expertise. Recall the physicist example; I may ask my physicist friend about the physical makeup of the table, and she replies with her expert judgment that the table is mostly empty space. This judgment is usually issued as a statement of fact; the expert makes no claims about its implications for action. With this in mind, let's compare three variations on Grand Master Hero Chess and see if there are any

differences in your reasons.²

First is Eavesdropping Grand Master Hero Chess. Recall that in Hero Chess you are playing with the fate of the entire world in balance. Instead of having the Grand Master as an advisor, though, in the Eavesdropping case the Grand Master is simply walking by the door where you are playing Hero Chess. You know this person is a Grand Master and as she walks by you hear her make a chess utterance that is relevant to the position you find yourself on the chessboard. This gives you some evidence in favor of believing that you should do as the Grand Master uttered.

Second is Hypothetical Advice Grand Master Hero Chess. In this case the Grand Master is available as an advisor; you can call her and ask how, if you were playing some random chess game, you should act in a particular circumstance. The Grand Master doesn't know the stakes of the game, just that you're playing chess. She gives you advice about how to move and this gives you some evidence in favor of believing that you should do as she advised.

Third is Authority Grand Master Hero Chess. In this case the Grand Master is in the room with you and knows the stakes. You don't ask for her advice but she commands you to make a certain move. Since she is an epistemic expert about chess, this command gives you evidence that you should make that move. (Given that practical reasons are veridical, whether you should actually make that move depends on whether that move will actually help you win the game.)

The question is whether the circumstances of the utterance make a difference to the

² To be clear, Grand Master is a title of practical chess experts, but we can safely assume Grand Masters are epistemic chess experts as well. Chess is a practical domain in which it is difficult (though certainly not impossible) to imagine a practical expert who also wouldn't be an epistemic expert.

reason the Grand Master's judgment constitutes for you. First, it is clear that in the Eavesdropping case, your evidence is weak and perhaps not even decisive. This is because the simple fact that an expert uttered a proposition with content in the field of her expertise does not mean that utterance is an expert judgment. Just before passing the door to the chess game, for example, the Grand Master could have prefaced her utterance with "If I were a bad player..." or "As a joke...". Because you are eavesdropping, you can't be sure whether the utterance is intended as an expert judgment. Normally we use contextual clues to set circumstances where it is understood that the expert is speaking qua expert and in the eavesdropping case you have none of those contextual clues. So even if you take her utterance to be some evidence in favor of believing that you should move in a certain way, it isn't decisive evidence.

The Hypothetical Advice and Authority cases are much closer, of course. In both cases it is understood that the Grand Master is making an utterance qua chess expert. In both cases, then, her judgment is decisive evidence in favor of believing that you should move in a certain way and your countervailing evidence, like your own judgment, is undercut. If there's no difference between these cases, it looks like evidential authority has the same effect on subjects' reasons as expert advice, and this would be problematic. Authority adds something extra.

The difference that matters is authority's intentional, illocutionary aspect, which I highlighted in chapter one. When an expert gives you advice, it may well have the effect of decisively changing your balance of evidence and undercutting countervailing evidence. But the expert needn't intend for this effect to happen. As an advisor, she intends her judgment to be a weighty piece of advice for you to judge, but it is still for you to judge. As

an authority, she must intend her judgment to not only be weighty but for it to be sufficient on its own, regardless of how you judge the balance of reasons to otherwise be. This is the illocutionary intent that Hobbes and Bentham identified.

Might not the expert intend for her advice to be overriding in the same way as the authority's command? She might, but then she has to express that intention in order to be an authority. Recall that what makes an utterance a command is not whether a particular imperatival locution was used but whether it was uttered with a particular imperatival illocution. This showed us that authority must be intentionally claimed; an authority can't accidentally command and thereby bind subjects. Subjects and their reasons must be the object of a particular intention. If an expert makes a judgment that she expects to be overriding but doesn't claim that it is overriding, it is rationally taken by the subject as a piece of advice rather than as a command.

To see why this matters, consider that in some circumstances it would be appropriate for the Grand Master to make different judgments in the Hypothetical Advice and Authority cases. The reason that she might make different judgments is that recognition of the practical context changes the threshold for decisive evidence, as I argued in the last chapter. Imagine that when you call her up in the Hypothetical Advice case, you are asking about a particular arrangement of play that makes it unclear what the best move is. The Grand Master weighs the options and comes to a 0.51 probability that you should make move R, so tells you to R.

Now imagine that as she tells you about R, you mention that if you lose the game, the world will be destroyed. This changes the case into Actual Advice Grand Master Hero Chess; she's no longer giving you advice about a hypothetical chess game but about this

chess game, against this opponent, with these stakes. When you are simply looking for hypothetical chess advice, she's comfortable advising you to R at a probability of 0.51. But now that she knows the stakes, she has to reassess; did she give the problem all the consideration it deserves? Is a probability of 0.51 sufficient to issue a judgment that she knows will decisively change your evidence in a game with the fate of the world on the line? Awareness of the practical stakes changes her role as expert advisor in important ways.

Similarly, from your perspective, you should be less confident in her advice when she doesn't know the situation. You justifiably worry that she did not consider her judgment to the extent that is necessary to be confident enough to act on it in this practical context. When she knows the context, you are more confident that her judgment was made in light of the high stakes, so should grant her judgment more weight. In both cases it may be true that you are confident enough in her judgment that you should completely defer to it, given her mastery and your relative inexperience. This point doesn't change the fact that the judgments differ in how they impact your balance of reasons in some circumstances.

This point is important but doesn't yet totally distinguish expert advice from authority. It is clear that in the Hypothetical Advice case you have to balance her advice with the fact that she doesn't know how high the stakes are, so her utterances have a different strength for you than in the Actual Advice case. But now we want to know whether Actual Advice differs from Authority in any significant respects: do her expert utterances, in full knowledge of the practical stakes, differ from utterances that she intends as commands to bind you? Yes, although again the difference is quite subtle.

When issuing expert advice, all that is necessarily implied is that the advice is

sufficient to give the layperson decisive reason to believe in some proposition. Even in practical contexts where it is clear that the advice has specific practical implications, advice only claims to change subjects' epistemic standing. When issuing a command with imperatival illocution, I claim to change your practical standing as well. (As argued throughout, this intent need not be to change what your duties are as long as it intends to change your practical standing in some way, as when the police officer intends to grant you a liberty to move along.) The is the difference between "You should believe that you ought to ϕ " and " $\phi!$ ".

When issuing a command as an evidential authority, something has been added to the claim that the judgment is sufficient reason to believe. This is a practical reason, but not in the right-making sense. Instead, evidential authorities intend to give subjects sufficient reason to *accept*.

Acceptance is a propositional attitude similar to pretense and importantly different from belief.³ Jurors, for example, have good reason to accept that the accused is innocent even if they believe she is guilty. They can perform their functions as jurors well only if they accept innocence because they are trying to come to a judgment supported by a certain restricted set of facts, those entered into legal evidence. Their belief in the accused's guilt may be justified by outside evidence, like how she has been treated in the media.

Whether you have a good reason to accept is a function of both your epistemic and practical contexts. It boils down to whether you have sufficient practical reason to act as if a particular proposition is true. Jurors have sufficient practical reason to act as if the

³ Michael Bratman, "Practical Reasoning and Acceptance in a Context", *Mind* 101 (1992): pp. 1-16.

accused is innocent, even if they are very confident she is guilty, because only if they act as if the accused is innocent can they secure their aim of justly evaluating the accused.

This illuminates the difference between the effect of the Grand Master's utterance in the Actual Advice and Authority cases. In Actual Advice, the utterance qua expert advice only claims to give you decisive epistemic reason. In Authority, the utterance qua evidential authority claims to give you decisive epistemic reason and, *in light of the practical context*, claims that your evidence is now *sufficiently* decisive to act as if it were true in these circumstances. Not only does your evidence change, it changes to the point that you should accept the proposition your evidence favors given the demands of the practical context.

Notice that your evidence can decisively change such that you shouldn't accept anything. Perhaps the Grand Master makes her utterance at such an early point in the game that there really is no fact of the matter about which move would be best. You might think there is good evidence that R is the best move at this point and the Grand Master's utterance could decisively change your evidence such that you rationally must believe that which move is best is indeterminate at that point. In that case it would be inappropriate for the Grand Master to issue a command even though she could perfectly well advise you that the reasons are indeterminate. Even though your evidence decisively changed by her utterance, it was evidence that didn't imply a reason to accept any particular course of action: it didn't entail an evidential duty. Issuing a command, on the other hand, implies not only that your evidence changes but that you have reason to accept the proposition in this practical context. It claims awareness not only of your epistemic reasons but your practical reasons as well.

Recall the case of Anecdote Mary, who gave her patient K even though her evidence that K would cure was merely anecdotal and so very weak, especially given that she had a pharmaceutical guide in her pocket. When Anecdote Mary gives K, she is culpable because she is negligent. Evaluating Anecdote Mary through the lens of acceptance is enlightening. The problem wasn't that Anecdote Mary gave K; so did Benevolent Mary and Data Mary, and neither was culpable for killing the patient. The problem was that Anecdote Mary's evidence that K would cure was clearly insufficient for her to accept that K would cure and so act as if K would cure. Given the morally weighty practical context, wherein the life of an innocent patient is at stake, her evidence was not sufficiently strong to give her a reason to accept that K would cure. Instead, because her evidence was merely anecdotal, she had reason to accept that she should check the pharmaceutical guide first.

It is only true that her evidence was insufficient in this particular practical context; if we change the particulars Anecdote Mary's culpability changes as well. This is because changing the details of the practical context changes the threshold at which her evidence is sufficient to give her reason to accept that K would cure. Consider Exigent Anecdote Mary: instead of a dermatological problem, the patient will die if he doesn't receive the cure within ten seconds. Exigent Anecdote Mary's practical context is very different; she doesn't have the time to check the guide, so should act as her evidence currently stands. The fact that it is weak evidence is still a problem. She may be culpable because she is a doctor who went to work knowing that certain pharmaceutical knowledge would affect the lives of her patients but without closely examining her pharmaceutical beliefs. Even if that's true, though, she's not as culpable as in the case where she has plenty of time to check her evidence and doesn't. When she has to act in the moment, her evidence can be sufficient to

generate a reason to accept even when it is much weaker than it would have to be in other contexts.

Exigent Anecdote Mary's case raises a concern. There it looks like the practical context is so demanding that less reliable evidence can still meet the threshold for culpability. In Hero Chess it's clear that disobeying the Grand Master is culpable and that not doing as your peer advised is nonculpable. But cases between those two extremes are less clear. The stakes are so high that it seems as if you should prefer any increase in the reliability of your chess playing, whether grounded in expertise or not. In Proficient Hero Chess, your advisor is another friend you beats you about eighty percent of the time. My intuition about this case is much less firm, but it still seems to me that ignoring your proficient friend's advice would appropriately bring the wrath of those whose lives depend on the outcome of your chess game. You would be culpable for ignoring her advice because she is an expert relative to you, even though she's not an expert more broadly speaking. Relative expertise can be sufficient to ground culpability because you can have sufficient reason to accept less than decisive evidence in demanding enough circumstances.

The expertise condition captures relative expertise perfectly well because it is concerned with the difference in reliability between two individuals. It does not require that the authority be expert in the most robust sense, an expert relative to all of humanity, or something of the sort. That said, for governments to achieve their aims they must have authority over most of their population, so it may be the case that more robust expertise is required. I discuss this concern more in chapter five.

Returning now to evidential modern political authority, the thought we have been exploring is that evidential authority requires expertise. The commands of an evidential

authority are distinct from expert judgments because commands add a practical claim, namely that the evidence is sufficient to constitute a reason to accept a proposition in a particular practical context. Call this *the acceptance condition* on evidential authority: the authority must not only claim to change subjects' evidence via her expertise but must also claim to give them sufficient reason to accept in a particular practical context.

The point isn't that reasons to accept are practical reasons in the right-making sense. Instead, reasons to accept are relevant to what makes people objectively vulnerable to punishment, as in Anecdote Mary's case. She's culpable because she accepted that K would cure when her evidence didn't support accepting that proposition, while Data Mary's evidence was sufficient to constitute a reason to accept so she wasn't culpable. Similarly, Exigent Anecdote Mary was not (or less) culpable because her evidence was sufficient to constitute a reason to accept in her exigent practical context.

When an expert issues a practical directive, it claims to assess not only your epistemic reasons but your practical reasons. And this is a problem, for epistemic experts in some domain may not have expertise about your practical reasons in general. If this is true, it's unclear how they can be practical authorities. I take up this issue in the next section.

2. The Precedence Condition

Despite adding the acceptance condition to the expertise and deference conditions, we have not yet reached a full picture. In this section I go on to argue for *the precedence condition*. The precedence condition addresses a problem for evidential political authority. Political authority is a form of moral authority, in large part because it claims that

disobeying the law makes one morally vulnerable to punishment. If political authority is moral authority, though, it appears that political authority requires moral expertise. This would be an extremely worrying feature of my account. It looks implausible that, of all the agents in the world, political agents could plausibly claim moral expertise.

Consider Grand Master Moral Maze Chess, wherein the Grand Master fulfills the expertise condition, issues commands so fulfills the acceptance condition, and the stakes entail that the deference condition is met. If this were sufficient for authority, the Grand Master would have moral authority over you, as in Hero Chess. The twist is the Moral Maze: the moral stakes are high but what to do morally is very unclear. Winning and losing both have moral advantages and disadvantages. The problem is that the Grand Master is a chess expert, not a moral expert. She can help you with chess moves, but she doesn't have the expertise to make her way through the moral maze and see whether you need to win this game or whether you need to throw this game. Although her judgment in the chess domain is expert, trying to follow her directives doesn't reliably lead you to act morally better because of her moral ignorance. If she understood the moral reasons appropriately, trying to follow her directives would lead you aright and she would have moral authority over you.

This makes it plausible that the Grand Master in Moral Maze can't be a moral authority precisely because she's not a moral expert. If she were a moral expert in addition to being a chess expert, she would know under what circumstances she has decisive moral reasons to issue you directives out of her chess expertise and she would understand the import of those reasons. Only then would her directives morally bind you, but she requires moral expertise and we are back to the implausible claim that modern governments must

have moral expertise in order to have genuine authority. The problem is that non-moral expertise is not sufficient to understand the moral situation, but understanding the moral situation is what's required for moral authority.

Another way to put this worry is to point out that the Grand Master does not require an understanding of relations between domains of reasons, only an understanding of the domain she is an expert with respect to, but understanding the relations between domains is precisely what is necessary to actually act well in the world. I can't act only on my chess reasons and leave moral reasons for another time. When the Hypothetical Advice Grand Master tells me her judgment, she is making that judgment purely on chess reasons and not in the context of my moral reasons, which is why the reasons she gives you to believe are weaker than the reasons given by the Actual Advice Grand Master.

Raz expresses a similar concern while discussing John and Ruth, experts on Chinese cooking and the stock exchange respectively:

Here the normal justification thesis establishes the credentials of John and Ruth as authorities in their fields. But whether or not there is complete justification for me to regard their advice or instructions as guides to my conduct in the way I regard a binding authoritative directive depends on my other goals. In such cases while talking of a person as being an authority one refrains from talking of him as in authority over oneself, and avoids regarding his advice or instructions as binding, even when, given one's goals, one ought to treat it in exactly the same way as one treats a binding authoritative directive.⁴

John and Ruth are experts in their domains, but they aren't experts about "my other goals".

Raz seems to be saying that they aren't authorities because they don't know about the other domains and the relations between those domains and their domain. I may be at the cooking class because I am collecting information on rival businesses and not to get better

⁴Raz, MF, pp. 64-5.

at Chinese cooking. If I have to make a worse meal in order to gain more information, my information gathering reasons are weightier than my cooking reasons, so John's cooking expertise can't ground authority over me.

There are two worries in the area. The first is that authorities only truly bind if the reason they give is "complete", i.e. all-things-considered. John and Ruth don't have authority because, while their commands bind within their domain, their domains are not important and often get overridden by reasons from other domains. This line of thought leads to the conclusion that the only genuine authority is authority whose commands always bind all-things-considered. On some views of morality, moral reasons are lexically ordered above all other reasons, so a moral authority would have genuine authority of this sort. But I don't think this is Raz's concern. As I explained in chapter one, domain-restricted authority is a clear and sensible notion. We know what it is to be bound within some domain, to get conclusive reason from a perspective without getting all-things-considered reason.

Further, this line of thinking is subject to a *reductio ad absurdum*, namely the claim that there can only be *global* moral authorities and advisors. That is, there is nobody who is a good moral advisor about one aspect of morality but not others. But this is false. Some people know more about racial discrimination because they care about it a great deal, have been the subject of it themselves, have spent considerable time and effort investigating it, and so on. It makes sense to go to them for advice about practical questions that potentially involve racial discrimination, but for exactly parallel reasons it makes sense to go to others for advice about gender discrimination. If we push the idea that moral authorities must be experts about all practical requirements we are forced to deny this

kind of domain-specific moral advice. For just like John wouldn't know when reasons from the cooking domain, the expert on racial discrimination wouldn't know when racial discrimination reasons are overridden by reasons from other moral domains, as when claims of racial discrimination run against claims of distributive desert. She would then give advice that would not reliably lead subjects to significantly better conformity with their preexisting reasons. Her advice wouldn't be good advice because it would often mislead.

John and Ruth can be authorities over us, and can bind us with their commands, despite the fact that their domains are restricted to cooking and finance and often the reasons they give will be overridden by reasons from other domains. Imagine I am in a situation where my moral duty is to ϕ while my legal duty is to ψ , and assume that, in this case at least, the moral domain overrides the legal domain, meaning that all-things-considered, I have decisive reason to ϕ . If I ϕ , I act rightly but illegally. Now a legal authority enters and commands me to not- ψ . ψ and not- ψ both entail not- ϕ , so I still act wrongly if I act legally. I still have all-things-considered reason to ϕ . But the fact that my all-things-considered moral duty did not change does not mean the authority did not have genuine *legal* authority over me. This is shown by the fact that her command changed my legal standing such that I act illegally if I ψ , whereas before the command I only act illegally if I not- ψ .

The intuition that she lacks authority over me is understandable, because she didn't change what I should do all-things-considered. But because domain-restricted authority is coherent, we can see what is wrong with this intuition. Her legal authority changed my legal standing, but that change didn't matter because my legal duties are overridden by my

moral duties in both cases. Had the legal authority *also* had moral authority, then she would have changed my all-things-considered standing. But since she only has legal authority, she only changes my legal standing.

If I go into John's Chinese cooking class with the intention of getting better at cooking, then (*ceteris paribus*) John's authority binds me all-things-considered.⁵ The strongest reasons I have for acting, here and now at the cooking class, are reasons from within John's domain of expertise. What this means is that John's domain takes *precedence* for me in this practical context. John's commands will give me decisive, domain-specific reason to act, and given that in this context John's domain takes precedence, I will often have all-thing-considered reason to do as John says. This doesn't mean that I must follow John's commands no matter what; a call from the hospital that my partner is deathly ill gives me a moral reason that overrides any cooking reasons I had to stay in the class, including any reasons given by John's commands.

This, then, is the precedence condition: if A has expertise in domain D and B does not, A can have authority over B in domain E only if D takes precedence in E (or D = E). So John can have moral authority over me because he has cooking expertise only when cooking is the morally precedent domain. If instead of Hero Chess I am engaged in Hero Cooking, such that unless I cook the best dish in the class the world will be destroyed, John has moral authority over me because of his cooking expertise. John can also have cooking authority over me when it is clear that cooking is the most precedent domain, as when I attend the cooking class in normal circumstances.

⁵ Given the nature of the domain, this is a very different sort of authority. First, it has no clear moral implications; disobeying doesn't mean much, although perhaps John appropriately gives me less attention or eventually kicks me out of the class. Second, a domain like cooking is particularly likely to fail the deference condition and so undermine authority.

But perhaps Raz's worry about John and Ruth is that they don't have authority over us because they don't *know* what our other goals are. John and Ruth don't know when acting on the basis of their expertise is practically required for me; their judgments are more like the Grand Master's in the Moral Maze. One way out of this problem would be to appeal to moral expertise. If John were a moral expert in addition to being a cooking expert, then he would know both when cooking takes moral precedence and what the best cooking reasons are. This could ground moral authority for John, but this again leads to the implausible position that modern governments must have moral expertise.

But moral expertise is not the only way out of this problem. People can know that a particular domain takes moral precedence without being moral experts. This is most obviously true in cases like Authority Grand Master Hero Chess. The Grand Master doesn't need to be a moral expert in order to understand that losing the game would be morally catastrophic, so she can know that her domain of expertise takes moral precedence for me. The precedence condition must be fulfilled for genuine authority, but knowledge of the precedence condition is also required. This knowledge can be achieved in a variety of ways, like when context makes the moral demands obvious. Moral expertise is not required for moral authority.

There is considerable intuitive resistance to the idea that experts on non-moral domains can have moral authority. I just explained why I think this resistance is mistaken, but this case is bolstered with an error theory for our intuitions in this domain. First note that almost all our interactions with experts will be outside moral contexts, and so almost always they are not moral authorities over us when we interact with them. When they issue judgments they may bind us with epistemic authority, making it irrational or

otherwise epistemically criticizable to believe anything other than what they say in their domain of expertise. Even on the rare occasion where they may issue practical directives, they will be overridden because their domains are not as important as other practical domains. We tend to think of authority in its most extreme forms, like god issuing moral commands that bind all-things-considered, meaning that we often don't recognize when we are bound by authorities whose commands are often, even regularly or mostly, overridden all-things-considered.

Further, sometimes we are morally bound to follow expert advice in a way that is very similar to being commanded by an authority, even when that expert doesn't have authority. The best example of this would be when an expert fulfills all the conditions for genuine authority but the acceptance condition. There are strong moral reasons for you to maximize, fulfilling the deference condition, their expert judgment reliably leads to significantly better conformity, fulfilling the expertise condition, and their domain takes precedence. You would act culpably if you ignored their expert advice. But they can't bind you morally because they don't intend to bind you morally, so they are not an authority even though you should treat their advice exactly as you would treat a command from them. Distinguishing authority from expertise from the perspective of the subject is thus often a very difficult, if not impossible, matter.

The pragmatics of claiming authority and giving advice in the cultural context where we generally pay advisors also complicates the picture. Consider a case where what you morally ought to do is completely determined by financial reasons.⁶ Let's say I go into the advisor's office knowing that my all-things-considered moral duty is to maximize my

⁶ This example is a modified form of one of Darwall's cases in "Authority and Reasons: Exclusionary and Second-Personal".

financial outlook this year. The advisor tells me “If you want to make the most money you can this year, you should ϕ .” This looks like advice, and there are plenty of reasons to couch her judgments in terms of advice. Most of the time when I consult a financial advisor she isn’t an authority over me, and as such, and given her position as a paid consultant, it would be inappropriate for her to command me. This inappropriateness carries through to cases where, as a matter of moral fact, she is an authority over me.

Further, the conditions that make an expert an authority are often going to be mysterious to a financial advisor. In general, acting in the real world is a complicated affair that implicates reasons from a host of different domains, so having one’s act completely determined by reasons from a single domain, and knowing that, will be quite rare. Thus a policy of acting as a mere advisor seems a good one for a financial advisor to take, even if sometimes her advisees are bound to obey her directives and she could have authority over them. She usually doesn’t have the time, resources, or need to determine whether the precedence condition is met. These considerations show why advisors often don’t issue commands, and perhaps shouldn’t, even when the conditions are such that they are genuine authorities. It is therefore no surprise that we feel experts aren’t authorities; in the rare cases where they are most authoritative, they still have good reasons not to act like authorities and issue commands.

Experts in the real world are very rarely moral authorities. This may have already been clear due to the implausible nature of some of the examples that I’ve been appealing to, like Hero Chess. Given this rarity, it is no surprise that we have considerable intuitive resistance to expertise grounding authority. This is especially true given that it is very difficult to tell when the conditions on genuine authority are met even when expertise is

clearly established: whether the deference condition is met depends on whether I have reasons to act on my own judgment for its own sake; whether the trustworthiness condition (below) is met depends on some understanding of the purported authority's character and knowledge; whether the precedence condition is met depends on common knowledge of moral reasons. None of these is easy for philosophers to determine, let alone the average layperson. But this shouldn't undermine our confidence in the claim that, when these conditions are met, an agent has genuine moral authority over another primarily due to her expertise.

3. The Trustworthiness Condition

The precedence condition addresses the worry that moral authorities must have moral expertise. But it might also be the case that moral authorities require moral *virtue*: in addition to knowing the moral context, they have to be good people who are motivated to act morally well. This is most obvious in cases where an expert is vicious so issues misleading directives.

Consider Vicious Grand Master Hero Chess. In this case she understands the moral reasons but she doesn't want you to win, she wants to see the world burn. So although the chess reasons indicate that you should R, she directs you to B instead. Are you bound to R?

You are not bound to R because the Vicious Grand Master does not have genuine moral authority over you. The problem isn't her expertise, the problem is her performance: she fails *the trustworthiness condition* on authority. The trustworthiness condition arises out of the performative nature of authority. Consider again the expertise condition: what matters is whether the subject is more likely to act better by trying to act

on the *directives* of the authority. Directives are importantly different than judgment. The Vicious Grand Master has expert judgment but doesn't issue expert directives. Her judgment is a function of her internal states. If you asked her hypotheticals like "If you wanted to win a chess match in such and such circumstances, how would you move?", she can give you extremely reliable, expert answers. But when she issues directives to you, i.e. when she performatively attempts to bind you with an utterance, trying to follow those directives will not lead you to better conformity with your preexisting reasons because she regularly misleads you out of her viciousness. Her failing is a failing of her moral character; she recognizes all the reasons but just doesn't care that issuing misleading directives to you will result in the world being destroyed.

As many theorists have emphasized, there is something troubling about deferring to either an authority or an expert. The problem is that when we defer, we do so without knowing all the relevant reasons. Consider Hardwig's characterization of the relationship between the layperson (A) and the expert (B):

So, if A accepts p on B's say-so, those reasons (B's reasons) which are necessary to justify A's belief are reasons which A does not have. Sometimes it is feasible for B to share with A all the evidence necessary to justify the claim that p. But usually not. Indeed, if A and B come from different disciplines or even different specialties within the same discipline, A often will not know what B's reasons are, much less why they are good reasons for believing p. // Thus, the blindness of A's knowledge that p: those reasons which are necessary to justify p (and A's belief that p) are reasons which A does not have. Obviously, since she lacks part of the evidence that justifies the claim that p, A is limited in the extent to which she can effectively scrutinize or challenge B's claim about p.⁷

Note that the layperson's knowledge is "blind". Raz says similar things about deferring to

⁷ Hardwig, "The Role of Trust in Knowledge", p. 699.

an authority.⁸ Deference seems to amount to putting blind faith in another person.

Deference is prima facie worrying, both rationally and morally, and deference to mistaken commands more worrying still. Morally, the ideal of the autonomous moral agent traces its origin at least back to Kant. For Kant, the moral agent must be autonomous or self-governed, only following goals she sets for herself. Deference to others seems to directly conflict with this ideal. Wolff complains that a subject “by refusing to engage in moral deliberation, by accepting as final the commands of the others... forfeits his autonomy.”⁹

Deference is rationally worrying because it requires a rational agent to act against what she judges to be the balance of reasons. You may think that the balance of reasons favors you continuing along your way in order to make it to work on time, but when the police officer issues you a command to stop, she purports to override that judgment. You are now required to act in a way you believe isn't supported by the reasons. The more important the subject matter, the more deference looks irrational, as when a soldier is commanded to go on a suicide mission or laws command citizens to act in ways they believe to be grossly immoral.

The reason that deference is sometimes not objectionable is trust. As I emphasized in chapter one, authority is importantly and necessarily relational. Like expertise, it involves one person making utterances that demand deference from another. So we should expect that the solution to the problem of deference has to do with the relationship between the two individuals. Whether we are in the epistemic or practical context, the only way to bridge the gap between individuals is with trust. Deference to another makes sense

⁸ Raz, AL, p. 24.

⁹ Wolff, “The Conflict Between Authority and Autonomy”, p. 27.

only if she is trustworthy.

To see how trust can make deference prima facie unproblematic, we need a more precise idea of trust. To my mind the best account considers trust to be special kind of reliance: you rely on the goodwill and competence of another by making yourself vulnerable to her in that you grant her discretionary powers over some interests you have.¹⁰ Trust is to be distinguished from mere reliance, which is simply a justified expectation that the object of the reliance will predictably perform some function. Thus you can rely on a sociopath to always act in their self-interest and you can rely on a thermostat to regulate the temperature. If either fails to act in the expected way, you may feel disappointment or confusion, but you do not feel betrayed, which is the characteristic response to a violation of trust.

Trust must have an agent as its object: you cannot trust your thermostat.¹¹ Intuitively, trust must involve more than a simple calculation of likely outcomes. If I can do all the probability calculations and so discover that you will perform as expected ninety percent of the time, and this in turn justifies my reliance on you, it does not look like I trust you at all. Instead, I merely rely on your calculated likely performance and the more robust claim of trust seems out of place. I am essentially treating you as a reliably performing machine. The extra element that trust adds, and the extra element that means only agents can be the appropriate objects of trust, is discretionary powers. When I trust you, I give you leeway to use your judgment to advance the interests I have entrusted you with as you

¹⁰ Annette Baier, "Trust and Antitrust", *Ethics* 96 (1986): pp. 231-60. The interests you entrust may be the interests of something else you have interest in, for example you allow the babysitter discretionary powers over some of your child's welfare and interests, which are also your interests.

¹¹ In terms of ordinary language, we sometimes loosely use 'trust' to mean mere reliance, a simple expectation of a possible outcome. But I am not in the business of ordinary language analysis, and this use of trust seems derivative of the more central notion I am explicating which requires an agent as its object.

see fit. Thus my trust depends on the exercise of your will, and must have an agent as its object. To trust you is to treat you as a person, not as a machine.

This is why trust also depends on the goodwill of the trusted; given the fact that the trusted has discretionary powers over the entrusted valuables, there will be plenty of opportunities for her to intentionally and subtly undermine the interests of the entrusted valuables. Giving her this opportunity is a necessary part of trust and is what distinguishes trust from mere reliance, but it is also what requires depending on the goodwill of the trusted.¹² Necessarily relying on the goodwill and judgment of the trusted is also what makes the truster vulnerable to the trusted. On this understanding, in order to trust another you must believe in her *goodwill* and her *competence*. Both of these are required because you grant her discretionary powers over the interests of something you value.

An illustration will help make this clear: when you get a babysitter for your infant, you trust the babysitter. You give him (limited) discretionary powers over the wellbeing of your child. If he is disastrously incompetent, and you know him to be so, then even if he tries his best to protect and advance your child's interests, he is going to fail. But because you granted him discretionary powers, those failures will result in harm to your child's wellbeing. Under those conditions he is untrustworthy and you shouldn't leave your child under his care.

Similarly, if he has malicious will toward your child, then he can use his discretionary powers to harm your child. He can do this very subtly, even within the discretionary powers you grant him. He may not kill your child, or take her to a smoky bar, neither of which you granted him power to do, but he may keep the child up much later

¹² This answers some of the concerns about Baier's account raised by Karen Jones, "Trust as an Affective Attitude", *Ethics* 107 (1996): pp. 17-8.

than is good for her, or feed her food that isn't the best. In both cases he has permission to exercise his discretion, over her bedtime or her food, but he does so with a malicious will in a way that is intended not to bring out the best results.¹³ He is not trustworthy. So in order to justifiably defer to another, the other must be sufficiently trustworthy: given the interests and discretionary powers you are entrusting she must be sufficiently competent and have sufficient goodwill. Only under these conditions will the trustworthiness condition be met.

The problem with Vicious Grand Master is not that she is not a moral saint but that she is untrustworthy. She's competent but doesn't have goodwill, so regularly issues misleading commands. Her failure is not of her expertise but of her performance and her relationship to you. Note that individuals don't have to be morally perfect, or even generally virtuous, in order to be trustworthy. A variety of mechanisms can establish competence and goodwill, although good moral character certainly helps with the latter. The important point is that experts sometimes fail to have moral authority because they aren't trustworthy, but we don't need to go on to claim that moral authorities must have all the virtues or anything of the sort.

4. The Expertise and Trustworthiness of Political Agents

We now have the following five conditions on genuine authority: the expertise condition, the deference condition, the acceptance condition, the trustworthiness condition, and the precedence condition. Fulfilling these conditions gives an agent genuine evidential

¹³ Parents who leave a list of precise activities and times for the babysitter trust their babysitter less than if they didn't leave such a list: they are granting the babysitter discretion over fewer domains.

authority over another, but they are not the only way to attain genuine authority.¹⁴ They are jointly sufficient but not necessary because other routes can secure authority, like actual consent.

In building the case for evidential governmentalism, it must be shown that at least one modern government meets these conditions. In the next chapter I take up particular cases, but in the remainder of this chapter I want to lay the groundwork for the more applied investigation by establishing the *prima facie* plausibility of the claim that modern governments could even possibly meet the expertise and trustworthiness conditions. The other conditions are either easily met or are a function of subjects' preexisting reasons and so not under the control of the government.

The expertise and trustworthiness conditions depend on characteristics of the purported authority and both conditions are places where it might seem modern governments are particularly apt to fail. We generally see modern political agents as bureaucrats embedded in massively complex, inefficient institutions. Our paradigmatic image of the expert is the brilliant individual scientist while the paradigmatic image of the political agent is either the overworked, apathetic bureaucrat or the self-serving, hypocritical politician. This characterization of political agents generates doubts about whether modern governments can meet the expertise and trustworthiness conditions.

First, we might be worried that political agents are simply too incompetent to be experts. We could spell this out in a number of ways. Politicians are chosen for their looks, their media campaigns, their ridiculous promises, but certainly not for their expertise! And bureaucrats are the paradigm of incompetence, any efforts stalled and vitiated by

¹⁴ Just like Raz's normal justification thesis, which is "normal" because it is the usual but not unique way to attain genuine authority.

labyrinthine institutional requirements, and any good intentions smothered and dissipated.

To answer this concern it is important to see that the commands of modern political authorities are generally the result of combining the input of many individuals. In large part political authorities can be competent because they are institutional authorities that can aggregate expertise. If I am asked to make some judgment about speed limits, in order to understand all the relevant underlying reasons I need to be an expert about, among other things, the metals and other materials that make up cars and their response to stressors, the biomechanics and physical vulnerabilities of individuals during accidents, the materials that make up roads, the tendencies of drivers at various speeds, the type and frequency of precipitation or other complicating factors, and so on. To do this responsibly requires aggregating the expertise of many different fields.

No individual could possibly be sufficiently competent in all these areas on her own, and to expect that only individuals can be experts comes from an overly individualistic understanding of our epistemic practices. As Hardwig emphasizes, even in the hard sciences, paragons of epistemic practice, it is groups that most commonly reach expert judgments.¹⁵ He looks at a published paper in particle physics that included nearly one hundred individual authors. No single author had the expertise to evaluate all the claims in the paper, and none was expected to. Instead, they aggregated their expertise and came to a single conclusion. This is a very common practice, and scientific progress as we know it would be impossible without the possibility of aggregating expertise. Good institutional structures can aggregate expertise and thus reach heights of competence that individuals cannot.

¹⁵ John Hardwig, "Epistemic Dependence", pp. 346ff.

This is why we needn't expect the politician in the legislature or the bureaucrat manning the desk at the front of the DMV to be experts. Governments employ huge numbers of people, including people with expertise, like scientists. The expertise lies with the scientists as individuals but is aggregated, with the expertise of others, in institutional structures. The "outputs" of these institutions can thus be appropriately considered expert judgments even though the leaders of those institutions may not be experts in the relevant fields themselves.

Perhaps, though, this is the very source of the concern. In the case of the physics paper, all the participants were experts in the general field. We would expect some of them to be able to aggregate the expertise of their colleagues, even if they didn't know quite as much as the others. But when political appointees or elected officials head political institutions, we have no reason to expect them to be able to *successfully* aggregate the expert judgments; that is, not only bring them together but weigh them against each other in such a way as to "retain" the expertise. If we hand over the implementation of expert judgments to fools, we will get foolish results.

The answer to this problem lies in institutional design. The answer is not that all political institutions will be expert, even when they employ experts. Bad institutions will undermine the aggregation of expertise, just as they will undermine trustworthiness. The aggregation of the physicists' expertises was done informally, but we trust it was done well because it was done by the experts' peers in some significant sense. The expertise of the aggregator seems to convey validity on the aggregation. If the aggregation of expertise in political institutions happens informally like this, we have good reason to doubt its validity. But institutions have the great advantage of being able to set explicit and public standards

of aggregation.

This is why it shouldn't matter (up to a point) what particular individual holds a particular bureaucratic office. The office functions as it does in large part because of the institutional constraints, not just because of the individual who holds it. We shouldn't deny the importance of individuals; all institutional rules can be gamed, distorted and repurposed. The institutional culture still matters a great deal. But institutional rules go a long way explaining how non-experts can sometimes successfully aggregate expertise, and so how institutions can have expertise that exceeds the expertise of any individual.

This point is very important. When evaluating the expertise of a particular political agent, we want to know whether the *directives* are expert. It may be the case that many individuals in the institutional structure lack the appropriate expertise. This is not a problem as long as the directives of the political agent are themselves expert. One way of putting this is that expertise is an emergent property, due not to the expertise of a particular individual in the institution but due to the combination of individuals and institutional constraints. Without appropriate institutional constraints, the expertise of the individuals won't be translated into expert directives. It is only in the context of appropriate institutional constraints that we can say the directives of a modern government are expert. This isn't mysterious, and my claim isn't that a real collective agent with expertise emerges from the institutional structure. I don't think this claim is necessarily false, but it is not required to understand how expertise could be an emergent property of individuals working within a particular institutional structure.¹⁶

The other worry is about trustworthiness. Just as we might expect bureaucrats and

¹⁶ For a recent case in favor of the "reality" of collective agents, see Christian List and Philip Petit, *Group Agency: The Possibility, Design, and Status of Corporate Agents* (Oxford: Oxford UP, 2011).

politicians to be incompetent, we might expect them to be untrustworthy. If political agents are radically untrustworthy, their expertise won't matter and they won't be authorities. They fail the trustworthiness condition and so lack authority because untrustworthiness gives subjects a good reason to reject the purported authority's directives. Experts whose expert judgments are rationally unavailable to us cannot be authorities.

Recall the account of trustworthiness I sketched above: there are two aspects of trustworthiness, competence and goodwill. An agent is trustworthy to the extent that she can justifiably be given discretionary powers over valuable interests. The extent to which an agent must be trustworthy depends on the entrusted interests; the more valuable the interests, and the more discretionary powers given over them, the more trust required and thus the more trustworthiness required for rational deference. Another way to think about what degree of trustworthiness is required is in terms of vulnerability: the more vulnerable I make myself to you by trusting you, the more trustworthy you must be for that trust to be justified. How vulnerable I am to you depends on the value of the interests I entrust to you, and how much discretionary power I give you over them, so depends on the same considerations.

These points are important because we grant modern political authorities nearly *complete* discretion over some of the *most important* interests in our lives, like our collective security and wellbeing. As the twentieth century demonstrated again and again, being subject to untrustworthy political agents, whether incompetent or malicious, can be epically disastrous. It appears, then, that political authorities require huge amounts of trust from their subjects, but that this degree of trust is the kind of thing political authorities are

particularly inapt to attain.

The question for political authorities is how they can establish goodwill and competence sufficient to make them trustworthy. The relationship between goodwill, competence, and trust is a complicated one. It is clear that complete competence cannot ground trust if it is matched with malice. Similarly, complete goodwill cannot ground trust if it is matched with complete incompetence. There needs to be both sufficient goodwill and sufficient competence. But it seems to me that beyond the extremes, there are some tradeoffs between goodwill and competence that make sense. I might not be convinced my brother is the most competent person in the world, but I know he loves me and always considers my interests in his decisions; he has goodwill towards me. It is possible that I can trust him just as much as I trust my financial advisor, who is cold-hearted and does not evince goodwill for me, yet is exceptionally competent in her domain (the only domain in which I trust her). It appears that sometimes a lack in competence can be made up for by goodwill, and some a in goodwill made up for by competence.

This is good news for political authorities, because there are significant limitations on their goodwill. They need to secure a sufficient level of goodwill, but they will make up for limitations on goodwill with competence. As I argued above, the institutional aggregation of expertise can, when done well, result in judgments that are ultimately more expert than the judgment of any individual would be. If we can establish some minimal level of goodwill, this advantage in competence will explain the possibility of modern political authorities being sufficiently trustworthy. Other than the common picture of government and bureaucracy as inherently untrustworthy, we have good reason to suspect that there is a limit to how much goodwill we can ascribe to a political authority. I take the

limit case of goodwill to be something like the brother case, or a partner, where you know the other person loves you. Not only do they take your interests into account, they often sacrifice their own interests for yours. This is the level of goodwill that seems required to justify sufficient trustworthiness to grant someone, like a partner, extensive discretionary powers over extremely valuable interests, like co-ownership of a house, a bank account, or raising a child.

This kind of goodwill can only be established by intense and prolonged personal relationships; the problem for modern political authorities is that there is no way all representatives of political institutions could have this kind of relationship with all those subject to their authority.¹⁷ As noted in chapter two, one of the defining and yet problematic features of modern political authority is their scale. You would have to personally know your mailperson, your state representative, your city council members, your senators, the president, and many thousands more to have this kind of trust in the goodwill of political authorities. In turn, the city council members would have to personally know you and every other citizen of their city, the senators every citizen of their state, the president every citizen of their country. Clearly it is impossible for any single human being to establish and maintain that number of close, personal relationships. Modern political authorities cannot reach anything close the highest level of goodwill. Not only can we not have close personal relationships with our political authorities, but due to the large numbers of people in our political bodies, it is unavoidable that most individuals we interact with as representatives of our political authorities will be complete strangers,

¹⁷ Cf. Elizabeth Anderson, *The Imperative of Integration* (Princeton: Princeton UP, 2010): p. 106.

and the goodwill we attribute to them thus very limited.¹⁸ But this is worrying because as mentioned we grant political authorities extensive discretion over some of the most important things in our lives.

I again respond by appealing to the institutional nature of modern political authorities. When I meet a stranger, I simply lack experience with them and so lack knowledge about their character and their attitudes and dispositions towards me.¹⁹ Limited interactions with them will not help me gain much in the way of this kind of knowledge. And this is how we interact with almost all political authorities, and their delegates: when I go to the DMV, the bureaucrat behind the desk is a stranger to me, and whatever limited interactions I have with her will not give me reason to think she has goodwill towards me and will consider my interests in her decisions, even when there are costs to doing so. But when she is acting in her capacity as a bureaucrat, she is embedded in an institutional structure.

The institutional structure is crucial: first, it can force her to consider my interests even when she personally would not, and second, it can do so publicly. This is to say that institutions can have moral characters that differ from the character of the individuals occupying institutional roles, and that institutional moral character can be public (and so knowable to me) in a way that individual moral character cannot, or cannot easily.

Individual moral character is more whimsical than institutional moral character, as it can

¹⁸ This is perhaps one reason why the leaders whom demand the most trust, those that have the most discretion over the most powerful and valuable things like the president, are also subject to the most personal scrutiny. It matters a great deal more whether I can trust the president than whether I can trust the DMV worker. The president has much more discretion over much more important interests.

¹⁹ Some theories of trust and testimony posit that there is a baseline of trust we automatically grant to strangers. Even if true this issue is orthogonal because that baseline is much too low to establish trust about the sorts of important matters we are concerned with. The baseline, stranger trust allows us to trust stranger's judgments about streets in their hometown or the time of day, but not much more.

change more discreetly and without apparent cause. I do not have direct access to the minds of those around me, and so may justifiably be worried that something dramatic has changed without me knowing it. (Think of courtiers living in fear that the monarch's goodwill can swing in an instant, for no good reason, leading to their death.) Institutions more easily carry their character on their sleeves, with explicit institutional rules and norms constraining the actions of the individuals that occupy roles within the institution.

When institutional power is exercised by an individual as a representative of that institution, I do not need to depend on the goodwill of the individual. Instead, their goodwill is in large part secured by the institution. I know that even if the bureaucrat in front of me is a vicious, selfish person she is still constrained by institutional checks. I can expect her to act in certain acceptable ways, ways that include my interests and so express goodwill towards me, because I know about these external, institutionalized and codified checks on her behavior. The institutional constraints significantly bolster my trust in her compared to how much I would trust her as an independent individual and a stranger. Of course, I still have to trust her personally to a certain extent. Even the best institutional constraints can be bypassed and gamed with enough effort. But I do not have to trust her personally very much, and much less than if she was simply a private individual to whom I was considering deferring. Institutions can make up the lack of goodwill we often assume when meeting a complete stranger, so the fact that political authority in modern society will necessarily involve trusting many complete strangers does not block political authorities from possibly being trustworthy enough to justify deference to them.

The institutional nature of modern political authority is thus of central importance to my account. It explains how expertise can be aggregated and then attributed to the

decisions of the authorities and it explains how political institutions can be sufficiently trustworthy to rationally defer to them. Further, it explains why their expertise grants them the power to command across so many domains. When a single individual claims to know everything about a domain, that is when we are confronted with a single expert, there are many ways to doubt her claim to expertise. For example, we might suspect that there are many factors outside her ken which are relevant to the domain but which she could not have possibly accounted for. Further, any worries about her trustworthiness, for example questions about her moral character, will infect all of her judgments.

Institutions avoid these problems. The idea is simply that an institution regulating, for example, transportation need not have an individual with all the necessary skills and expertise. Instead, it employs experts in a wide variety of fields: metallurgy, public health, environmental impact, etc. If an individual comes to me claiming to be an authority about all things transportation, I have very good reason to doubt her. There's simply too much to know. But an institution that publicly employs a range of experts and has institutional mechanisms for combining those various expertises can have a plausible case for being an expert on a very wide range of issues, wide enough to cover the domain of transportation, and thus can be justifiably trusted as an authority over most people, in a broad and important domain. Their expertise and trustworthiness does not necessarily change when different individuals are in their institutional roles. Individuals die, retire, and so on, but the institution and the institutional features that ground competence and goodwill remain.²⁰

²⁰ This does not amount to the claim that you trust the institution. Trust, as noted, must have an agent as its object, unlike mere reliance. There are two salient possibilities: you are trusting the individual

Before concluding, it is interesting to note the relationship between markers of trustworthiness and other discussions of political authority. Accounts of political authority like mine often look sterile and overly moralistic. The focus is on conformity with right reasons, and all the usual values people appeal to when they explain their allegiance to political authorities are put aside. Other accounts of political authority do a better job on this front. They focus on particularistic ties, shared political cultures, and so on. What my analysis shows is that these factors do not directly justify political authority. But the standard that justifies political authority, conformity with right reason, cannot also be the standard of the decision procedure subjects use to identify and follow authorities. This is because laypeople lack the competence to evaluate expert judgments and because right reason is not identifiable qua right reason.

What this means is that the factors people should use to measure trustworthiness will often be those factors we commonly call upon: commitment to a certain set of political ideals, shared culture, and so on. These are markers of trustworthiness (some better than others) and markers of trustworthiness are all individual citizens have to base their decisions on. So it is often valuable to focus on developing shared identities, on encouraging the markers of trustworthiness. But we shouldn't mistake markers of trustworthiness for what ultimately justifies our trust in them, namely successfully helping us act better in the world.

In this chapter I have argued that genuine evidential modern political authority requires an agent meet three further conditions, the acceptance, trustworthiness, and precedence conditions. In combination with the expertise and personal conditions we have

representative of the government, but qua institutional role, or more controversially you are trusting the corporate agent, the government, that is constituted by individuals inhabiting an institutional structure.

jointly sufficient conditions on genuine modern political authority. I also made the case that institutional mechanisms make it prima facie plausible that modern governments could attain the required expertise and trustworthiness. Plausibility, though, is not enough to meet the philosophical anarchists' challenge. In the next, final chapter I argue that some modern governments actually meet all five conditions on genuine modern political authority and so that evidential governmentalism is true.

5 – Towards Evidential Governmentalism

We have the full theory of evidential modern political authority before us. The task now is to return to the question of political anarchism, and specifically we want to know if any modern governments plausibly meet the five conditions on genuine evidential modern political authority: 1) expertise, 2) deference, 3) acceptance, 4) trustworthiness, and 5) precedence. If so, then evidential governmentalism is established: at least some governments have evidential modern political authority over some of their subjects. If this is true, then philosophical anarchism is false.

This chapter makes the case for believing that some Western democracies have genuine evidential modern political authority over most of their subjects in most domains. I proceed by first considering what duty evidential authorities might plausibly be thought to be expert about. If we can find such a duty and it is important enough, the deference condition will be met. Further, if modern governments can plausibly attain expertise about this sort of duty, the expertise and precedence conditions will be met. I then argue that this duty covers large enough proportions of the population for the government to actually achieve its aims. Finally, I appeal to democracy to further bolster the case in favor of the government's expertise and argue that democratic procedures are an invaluable component in meeting the trustworthiness condition in the modern world. In the end, we will see that some modern governments have political authority such that subjects are objectively morally vulnerable to punishment for disobeying the law.

1. The Natural Duty of Justice

The notion of evidential political authority depends on a situation where there is some sufficiently important moral duty that binds subjects and that the authority has relevant expertise for. If there's no such duty, then the precedence and deference conditions won't be fulfilled and governments can't have authority. This is why the Grand Master has moral authority over you in the Hero Chess case but not in the Fun Chess case even though she has chess expertise in both cases. If we're traditionalists about political authority the duty is built in because traditional authority can create new moral duties, but the whole idea of evidential authority is that it cannot. So what duties are there to be expert about?

There are some obvious ones, like the duty not to murder, not to steal, and so forth. So one strategy to establish the wide-ranging authority of government would be to proceed in a piecemeal fashion, picking out all the duties of this sort and explaining why the government has expertise about each of them. But it's unlikely that most of the government's activities can be grounded in such obvious natural duties. What specific duties are we fulfilling when we pay taxes that are used for art museums, highways, and so forth? While appealing to natural duties could explain how *mala in se* laws make citizens vulnerable to punishment, it doesn't look like this strategy could justify *mala prohibita* laws precisely there's no preexisting moral duties that *mala prohibita* laws map on to. But an account of political authority that can only capture *mala in se* and basically says *mala prohibita* laws don't bind citizens would be insufficient.

Instead of trying to find a conjunction of piecemeal duties that could cover all government activities, governmentalsists have appealed to a variety of wider, more general duties. These include the duty of fair play, Samaritan duties, associative duties, and others.

I'll focus on one particularly prominent case, Rawls' natural duty of justice, which finds its roots in Kant.¹

Kant claims that all people have a natural duty to create and enter just political institutions. The reason for this is that respect for others requires us to have an overarching authority that can settle disputes, without which we subject each other to the constant threat of unjust harm. Not only does this natural duty of justice require setting up domestic institutions, it also requires states to enter into international institutions. For if they didn't, they would subject their neighbors to the unjust threat of harm, just as individuals do.

In this same vein, Rawls argues for a natural duty of justice. Rawls doesn't extend his natural duty to the international sphere as Kant does, but at the domestic level the demands are similar. Rawls writes,

[F]irst, we are to comply with and to do our share in just institutions when they exist and apply to us; and second, we are to assist in the establishment of just arrangements when they do not exist, at least when this can be done with little cost to ourselves. It follows that if the basic structure of society is just, or as just as it is reasonable to expect in the circumstances, everyone has a natural duty to do what is required of him. Each is bound irrespective of his voluntary acts, performative or otherwise.²

So citizens have a natural duty to obey the law because obeying the law is how we discharge our natural duty of justice in a just society. The natural duty applies to us regardless of consent, so could plausibly justify the actions of modern governments.

The main question is how the fact that we have a natural duty of justice easily translates to a universal duty to obey the law. In the quote, Rawls says that it simply

¹ Raz also appeals to Rawls' duty of justice, though for slightly different reasons. See Raz, MF, p. 66.

² John Rawls, *A Theory of Justice*, pp. 293-4.

follows that “everyone has a natural duty to do what is required of him”. The “required” is the important part. Rawls is a version of what Simmons calls a “necessity” theorist, who appeals to the necessity of obeying the law toward the fulfillment of a more general duty.³ The thought is that you have a duty to obey the law because if you didn’t obey the law, justice couldn’t be achieved. Since you have a duty of justice that can’t be met if you don’t obey the law, you have a duty to obey the law.

The problem with this argument, as Simmons points out, is that it’s simply false that if you didn’t obey the law, justice couldn’t be achieved.⁴ It’s true that if most people didn’t obey the law most of the time, justice couldn’t be achieved. But in many instances breaking the law has hardly any effect at all, let alone a catastrophic effect.⁵ When you speed, or jaywalk, or run that infamous desert red light, justice can still be achieved. Justice is resilient and doesn’t require perfect conformity to achieve (good thing, too).

However, if we look at the natural duty of justice as the grounds of evidential political authority rather than traditional political authority, the picture is much different. Instead of arguing that the natural duty of justice grounds a duty to obey the law, we want to establish the claim that the natural duty of justice grounds vulnerability to punishment. I’m still going to make a necessity argument, but it’s that justice requires equal vulnerability to punishment for disobeying the law rather than requiring universal obedience to the law.

No matter your theory of justice, one of the necessary elements is a fair system of punishment. Everyone agrees that justice requires a significant degree of obedience to the

³ Wellman and Simmons, *Is There a Duty to Obey the Law?*, pp. 127ff.

⁴ *Ibid.*, p. 167.

⁵ This also seems to be the main problem with one of Socrates’ reasons for not fleeing Athens prior to his death in *Crito*. *Crito* contains nascent forms of many modern accounts of political obligation.

law, just not universal obedience. Despite the fact that a fair system of punishment can never secure perfect compliance, without a fair system of punishment you couldn't get the significant level of obedience required for justice. Many people respond to the threat of punishment in the law rather than the normative imperative, and if that threat were removed they would act in many unjust ways. Not only would they violate others' negative rights more often, they also wouldn't contribute to the sorts of public projects that justice requires, like public education.

A fair system of punishment has two important features for our purposes. First, it must have the liberty to punish. If it didn't, then it would wrong citizens every time it punished them and would not serve the ends of justice. Notice that the liberty to punish correlates only to a vulnerability on the part of subjects, not a duty. The extension of this argument to a duty to obey the law is the part that Simmons objected to.⁶ The state is at liberty to punish people for not meeting their duties of justice for essentially the same reason that you are at liberty to break into someone else's cabin in the woods in order to survive. Conditions of moral necessity give individuals the permission to encroach on others in ways that would be objectionable without the moral necessity.

Second, a fair system of punishment must exercise its permission to punish in an effective and non-arbitrary manner.⁷ Imagine that the government issued a law that made group A vulnerable to punishment for ϕ -ing but didn't make group B vulnerable to

⁶ Although Simmons, as with many anarchists, also denies that the government is required, arguing that private institutions can meet any demands the government can. This seems quite implausible to me, but I'm not going to enter that debate here.

⁷ Why not think a fair system of punishment has to punish everyone equally, rather than just equal vulnerability? Well, like obeying the law, equal punishment is not required for the system to meet its ends. Equal vulnerability to the law is required in order for the threat of punishment to have its deterrent effects, which I take it are central to the aim of justified punishment. For more in this vein, see Quinn, "The Right to Threaten and the Right to Punish".

punishment for ϕ -ing. This is blatantly arbitrary and prefers B to A for no good reason. Arbitrariness, though, renders a system of punishment unfair, and justice cannot be achieved with an arbitrary system of punishment. A fair system of punishment has to treat each individual equally, which is one reason laws targeted at particular individuals are bad laws. It also has to apply broadly; if most people weren't vulnerable to punishment, it wouldn't be able to sufficiently coordinate action to achieve justice. So a fair system of punishment not only depends on the vulnerability to punishment of subjects but also to creating widespread and non-arbitrary vulnerability.

Widespread and non-arbitrary vulnerability to punishment is necessary for a fair system of punishment, and a fair system of punishment is necessary to achieve justice. The necessity argument here is not about establishing a duty to ϕ because ϕ -ing is necessary. Instead, the necessity argument shows why the natural duty of justice is important enough to make citizens vulnerable to punishment if they don't meet it as well as why vulnerability must be sufficiently widespread. Not all moral duties are of the sort that people can be punished for failing to meet, especially punished by the government, like the consent-based duty not to cheat on one's partner. Once we see that justice is a sufficiently important aim and that widespread, non-arbitrary vulnerability to punishment is necessary to achieve this aim, we see why citizens have no moral objection to being made vulnerable to punishment by the directives of the government. This also means that governments meet the deference condition, for fulfilling the duty of justice is an area where it is more important to complete the task than to make the decision on one's own. The duty of justice is doing the same work in establishing governments' evidential modern political authority as the duty not to let the world be destroyed did to establish the Grand Master's authority over you in Hero Chess.

One of Simmons' worries about necessity theories is that they are unclear about the nature of necessity and often make arguments that elide necessity and helpfulness. The problem for grounding a general duty to obey the law is that obeying the law on every occasion is merely helpful and not necessary to achieve justice. But the sense of necessity I am appealing to is very robust. My claim is that one of the constitutive elements of justice is a fair system of punishment and that one of the constitutive elements of a fair system of punishment is widespread and non-arbitrary vulnerability before the law. Thus, if people are not equally vulnerable, nonarbitrariness is entailed, and neither a fair system of punishment nor justice can be achieved. The necessity is not about causal contribution as a piece of a particular outcome but about what constitutes the outcome. This view also recognizes the fact that obeying the law in every case is not necessary to achieve justice; even if I'm vulnerable to punishment, I may decide that the risk of punishment is worth breaking the law. What would undermine justice is if I wasn't vulnerable to punishment simply because I decided not to be on this occasion.

Note that despite the fact that widespread and non-arbitrary vulnerability is required, this does not rule out some people not being vulnerable for good reasons. The government often tries to make these sorts of accommodations, as with religious exceptions to government restrictions on peyote. The vulnerability that's required for justice is equal application of the law to all subjects, including the ability of everyone to get exemptions where they meet the conditions; anyone who uses peyote for religious purposes can get an exemption. This kind of exception is neither arbitrary nor widespread enough to undermine a system of punishment being effective and fair, and is a feature of a good legal system.

This discussion of the natural duty of justice may look a bit orthogonal to my concern with political authority. Simmons and others are interested in the perspective of the citizen and whether she has a general obligation to obey the law. Given my rejection of traditional accounts of political authority, I have to agree with Simmons that citizens don't have such an obligation.⁸ Instead, for evidential modern political authority, I was trying to identify a duty that could plausibly be thought to cover all or most subjects and that is important enough that it can meet the precedence condition. The natural duty of justice meets both of these criteria; it is the general sort of duty that everyone has and it is of the sort that people can appropriately be vulnerable to punishment for failing to meet. So the natural duty of justice can ground evidential authority, but that is distinct from whether it can ground a general duty to obey the law.⁹

If the precedent duty relevant for evidential modern political authority is a Rawlsian duty of justice, then it follows that the expertise required to meet this duty is expertise about setting up, maintaining, and upholding just institutions, particularly the institutions that make up the basic structure of society. The question now is whether it is plausible that any modern governments have this expertise with respect to some of their citizens. And I think it is. In the last chapter I gave some schematic reasons to believe that governments' institutional nature could allow them to achieve heights of expertise that individuals could not. In order to establish the expertise condition, though, I need to make the case that actual governments have the specific expertise that is required to meet the duty of justice.

⁸ Although it's worth noting that my account doesn't entail that there's no duty to obey the law, it just claims that no duty to obey the law is required for the government to have genuine authority or achieve justice.

⁹ Cf. Wellman, "Liberalism, Samaritanism, and Political Legitimacy".

In order to give content to the duty of justice we need a theory of justice, so I am going to continue in the Rawlsian vein of this discussion and assume that the best theory of justice is Rawls' own justice as fairness. The content of the duty of justice is thus to create, support, and uphold institutions that realize Rawls' two principles of justice. Whether you agree with this assumption isn't particularly important, as the argument below can be made with your preferred theory of justice, *mutatis mutandis*, although applying a different theory of justice may result in different evaluations of actual governments.

Given that you are reading this dissertation, you are a particularly inapt example of the sort of citizen that governments might have justice expertise with respect to. If your profession is the study of justice, it is plausible that you have justice expertise, and likely to a greater extent than the government. Below I'll introduce some considerations in favor of believing that even those with justice expertise are subject to the government's authority in broad swaths of their daily lives. But in order to assess the plausibility of the claim that governments have sufficient justice expertise to ground authority over many or most of their citizens, you must abstract from your own case. The number of individuals with justice expertise is essentially trivial considered against the huge populations of modern states.

While abstracting from your own case as an individual, you shouldn't also abstract to the perspective of any citizen in any state. The governments with the most plausible claim to justice expertise are those governments that are closest to actually being just. We see their expertise realized in their institutions and outcomes. With respect to justice as fairness, the best case for justice expertise is for Western, liberal democracies, and especially the Scandinavian governments. On the other hand, there is not even a prima

facie case for the justice expertise of some actual regimes, of which North Korea is the best, most horrifying current example.¹⁰ The cases between North Korea and Western democracies are wide-ranging and more difficult. For now, I'm going to focus on Western democracies, given that they have the strongest claim to justice expertise.

In particular, I'll consider the case of the American government for three reasons. First, I am most familiar with American political history and institutions, allowing a more nuanced discussion. Second, since I am American, whether the American government has genuine authority is a particularly relevant question. Finally, if I can build a plausible case that the American government has genuine authority, I will have also made the case for much of Northern and Western Europe, as those governments have stronger claims to justice expertise than the American government.

Let's start with the claim that the American government has justice expertise with respect to the average American citizen. This claim could be established in a variety of ways. One is simply to appeal to American political history. Some of the basic elements of justice had to be imposed on American society from the federal government and especially the Supreme Court because the laws of the land, set by legislators elected by popular vote, would not respect those elements. The most obvious example is racial equality and the civil rights movement. This is not to say that individuals involved in that movement lacked justice expertise, like Martin Luther King, Jr. or Malcolm X (or whatever figure you prefer), nor is it to deny that those individuals and many average citizens played a large, leading, important role in changing the attitudes of other citizens and the government itself. It's

¹⁰ For a recent United Nations report detailing the widespread human rights atrocities committed by the government of the Democratic People's Republic of Korea and its leader Kim Jong-Un, see <http://www.ohchr.org/EN/HRBodies/HRC/CoIDPRK/Pages/ReportoftheCommissionofInquiryDPRK.aspx>. Accessed 2/23/14.

simply to say that most Americans did not believe racial equality was a fundamental issue of justice (or, if they believed it, their political judgments did not cohere with their beliefs). They were vulnerable to punishment if they disobeyed the civil rights laws because the government was expert about this issue and they were not. We see the same top-down dynamic in many cases of basic justice, as even today the Supreme Court strikes down laws discriminating against the LGBTQ members of our community.

We can also see the non-expertise of the average American citizen by looking at opinion polls about basic issues of justice. Nearly one third of Americans support making Christianity the national religion of the United States.¹¹ One quarter of Americans believe the government should do little or nothing about rising income inequality.¹² Despite rapidly changing attitudes, about forty percent of Americans still believe that homosexual marriages should not be recognized by the state.¹³ We could go on. These are basic issues of justice from the Rawlsian perspective: racial, gender, and sexual-orientation equality, the separation of church and state, and the devastating political effects of income equality. Any political institutions that do not address these issues are not just, yet significant percentages of Americans disagree.

Just so we're clear, let's consider the actual numbers instead of percentages. There are about 315 million Americans. That means about 100 million people in America think Christianity should be the state religion, about 75 million think the government should do little or nothing about rising income inequality, and about 125 million think homosexuals

¹¹ HuffPo/YouGov poll, April 2013, http://big.assets.huffingtonpost.com/toplines_churchstate_0403042013.pdf

¹² USA Today/Pew poll, January 2014, <http://www.usatoday.com/story/news/nation/2014/01/23/pew-poll-obama-wealth-gap-sotu/4777385/>

¹³ Pew poll, 2013, <http://www.people-press.org/files/legacy-pdf/3-20-13%20Gay%20Marriage%20Release.pdf>

should not have the legal right to marry. If they are so mistaken about such basic issues of justice, though, it's plausible to think that they misunderstand justice broadly. Thus it is plausible to think that the American government, even with its widespread and egregious failings and imperfections, has expertise in the domain of justice with respect to those individuals.

Simply by looking at the most obvious cases of citizens' opinions on basic issues of justice, we have made a plausible case that the American government has genuine evidential authority over a third of its citizens, or about 100 million people. The point is not that the American government always acts well, or even mostly acts well, just that there are at least 100 million people in America whom understand justice so poorly that they will act better if they try to follow the commands of the American government than if they tried to achieve justice by following their own judgment. This is, it seems to me, eminently plausible, as it should be to anyone familiar with the American public.

Note that this claim is not about how well members of the public treat others interpersonally, but about how they would set up institutions that must meet the two principles of justice. It's not a surprise that the average person's judgment about the massively complex question of how to set up the institutions that constitute the basic structure of society in a way that treats all fairly is bad. The more complex the question, the more we should expect that reliably good judgment requires intense, prolonged inquiry and training of the sort that grounds expertise. Political questions, especially questions of the institutions of the basic structure, are among the most complex questions we face.

Establishing authority over 100 million people is no small achievement and grounds at least a presumption in favor of thinking that the American government could have

sufficiently wide-sweeping authority to achieve justice. But that can't be the end of the argument. Given that I have now grounded the possibility of genuine modern political authority in the government's unique ability to help us meet the duty of justice, it needs to be the case that governments have authority over a significant enough portion of their population to achieve justice, and over enough domains. I don't know that we can pin down a certain percentage, but one-third is not enough. We need reasons to think that some modern governments have genuine authority over most of their citizens, in most domains, over most of their lives, to establish that they have sufficient authority to achieve justice.

You might note at this point that the claim that the American government has justice expertise in the sense that it is expert about Rawls' justice as fairness is quite implausible. That's correct; the American government has no expertise about Rawls' philosophy and doesn't claim to be constrained by the Rawlsian principles of justice. This isn't a problem because expertise does not require perfect, or even objectively accurate, beliefs. Consider that we know Newtonian physics to be false, but that doesn't undermine the claim that pre-relativity physicists were physics experts. They endorsed a view that was false but their beliefs were still constrained by significant evidence and worked in a large percentage of test cases. Their expertise is based in their ability to correctly apply plausible and well-supported theories to new situations, not in their ability to apply the correct theory.

It's likely that modern physicists also have a host of false beliefs based on false theories that, although accurate on the whole, are not strictly true. For example, if physicists find evidence debunking string theory, some physicists' beliefs will be shown to be false. This doesn't render those physicists non-expert about physics. A government can

have justice expertise even if it does not endorse the correct theory of justice. What exactly justice requires and whether particular theories of justice are true is also the subject of much more reasonable disagreement (even among experts) than whether relativistic physics is true, so we can't necessarily demand that governments endorse justice as fairness in order to be expert in the way we would demand that a physicist reject Newtonian physics in favor of relativistic physics in order to be expert.

So we needn't demand that modern governments endorse justice as fairness in order to evaluate whether their directives reasonably meet the demands of justice as fairness and whether the average citizen would better fulfill her duty of justice by following the government's directives than by following her own judgment. Indeed, the argument in chapter four actually showed that when the stakes are high enough, and I take it the achievement of justice is an extremely important duty, non-expert yet still significant differences in reliability can ground authority, as in Proficient Hero Chess. The authority needn't have expertise in the most robust sense of the idea, just relative expertise over citizens. Of course it would be better if governments explicitly endorsed the correct theory of justice (whatever it is), as that would make them much more likely to actually achieve justice. Not making that endorsement, though, doesn't doom the possibility of having genuine authority.

The government's expertise also explains how laws can function as standing orders. Recall that in chapter two I argued that one of the defining features of political authority is that is diachronic to an extreme extent, claiming the power to bind citizens over the entire course of their lives. This is puzzling, especially if we are focused on the sorts of examples I have appealed to, where authority is temporary and arises out of what are essentially

emergency situations. In order to achieve justice, though, the government needs diachronic authority. In chapter three I characterized the expert's judgment as having trumping reliability; in addition, it has *insurmountable reliability*. Insurmountable reliability explains how the government's authority could plausibly bind citizens over the course of their entire lives. A's reliability is insurmountable for B iff it is rational for B to treat her judgment as permanently less reliable than A's judgment because it necessarily is less reliable or because it is not irrational for B to permanently forego the only process that could make her judgment as or more reliable than A's. Let me explain.

The kind of deliberation or following your own judgment that expertise rules out is the kind that aims to come to a conclusion in better conformity with the preexisting reasons than the expert's judgment. If such a conclusion were reached, then the layperson would be justified in ignoring the command and following her own judgment again; her judgment would have surmounted the expert. We might call this, somewhat awkwardly, surmounting deliberation. Raz actually hints at something like this while explaining why exclusion does not cut off deliberation as such: "There is no reason to prevent a person in such circumstances from going through the arguments to amuse himself or as an exercise, etc., *so long as he does not trust his judgment enough to act on it.*"¹⁴ If we plausibly interpret trusting one's judgment as a claim about confidence in the reliability of one's judgment, then Raz is ruling out deliberation that could affect confidence about reliability and eventually justify disobedience. The claim that the expert's reliability is insurmountable for the layperson amounts to the claim that the layperson has no hope of successful

¹⁴Raz, PRN, p. 48. My emphasis.

surmounting deliberation, and thus is irrational if she engages in it (for the purpose of surmounting, although she can deliberate for self-improvement or other reasons).

So, to show that laypeople must permanently submit their judgment to an expert's we must show the layperson has no rational hope of successful surmounting deliberation. Recall the "sustained, prolonged, and systematic" inquiry that characterizes expertise. Consider the expert on physics. It's possible for me to understand the physicist and her reasons in some sense: I'm a reasonably intelligent person, well educated, using the same language, and so on. But consider that the physicist has a graduate degree, so she has spent years of her life gaining knowledge in the area, thinking about the issues, testing her beliefs, discussing the possibilities with others similarly knowledgeable and passionate, and so on. From my current knowledge and skill base, it would take years of work and study to get to a place where I could fully understand the physicist's decision-making process and rationale.

We're considering whether I, a layperson with a limited amount of time, effort, and interest to deliberate about the issue, could reasonably think that any amount of deliberation, fact-finding, and so on could improve the reliability of my judgment in this domain to the point of the expert. And spelling it out this closely makes it clear that I could not. That is, it would take a major life disruption, e.g. getting a degree in another field, for me to be able to understand the expert directly. This is not a realistic possibility for me, and so I know that it is not a realistic possibility that I could look at the extant evidence and deliberate more reliably than the expert. No matter the amount of time I spend thinking about it or amount of work I put into it I will remain less reliable, unless I am willing to

embrace the major life disruption. And even then I would have to cede my judgment in all those other areas I can't understand without more major life disruptions.

It is not irrational for me not to pursue expertise in every field because that is quite literally impossible for a human with limited cognitive capacities and a limited life span. Thus when a layperson encounters an expert, she is rational in permanently deferring to the judgment of the expert because she knows that any amount of deliberation (excepting major life disruptions) could not reasonably be expected to increase her reliability to the point of surmounting the expert's reliability. This explains how it's not only the case that many average citizens will be subject to the authority of the government at a particular time, but also how they are subject over the entire course of their lives.

If the government has expertise about meeting the demands of justice relative to much of its population, we can also explain how modern governments' claim to set their own jurisdictional limits might be justified with respect to a significant portion of the population. Governments don't only claim to bind you with respect to domains and acts that you think they have authority over. Governments go on to make the much more robust claim that they decide which domains they have authority in: they set their own jurisdictional limits.¹⁵ Laws can both extend and restrict the domains into which the government imposes its will.

This is a kind of meta-authority: authority not only with respect to the political domains but authority to determine which domains are political. If this is right, then what looks necessary to justify this meta-authority is meta-expertise: expertise not only about how to achieve justice in particular domains but expertise about which domains are the

¹⁵ Cf. Raz, MF, p. 77.

appropriate concerns of justice. Justice expertise of the sort I have been ascribing to modern governments, though, is precisely this sort of meta-expertise. It's authority not only about how to secure the safety of citizens but the authority to declare racial, gender, or sexual-orientation equality the appropriate concern of the government. Given the examples I listed above, it is plausible that this is the sort of authority modern governments have over vast swaths of their population, particularly the non-experts about justice.

Once we focus on authority that can set its own jurisdictional limits, we can see where the justice expertise of modern governments actually rests. It is not the case that politicians or bureaucrats have expertise about jurisdictional limits; see the constant encroachment, or attempted encroachment, of laws into non-political domains, like the sexual choices of consenting adults. Passing laws of that sort doesn't necessarily render a government unjust precisely because modern governments employ a separation of powers. When the legislature passes unjust laws of this sort in governments with a plausible claim to justice expertise, the courts strike them down.

So is my claim that judges have justice expertise? I don't think that's terribly implausible in many cases, especially at the highest levels. But I also don't need to make this claim. Recall from chapter four that we needn't identify individuals within the government that have expertise. Instead, we are concerned with whether the directives of the government are expert, which is best explained by appealing to expertise as an emergent property of individuals embedded within a particular institutional structure.

More important than individual justices' expertise, then, is the emergent expertise of the political institution as a whole. The conception of justice that judges appeal to when striking down particular unjust laws is the conception of justice contained in the

constitution. The constitution of a state describes the character of both the state and its citizens, and defines the relations between the two, albeit in quite general terms. But those general terms pick out the basic elements of justice, according to the state, that laws cannot impede upon. A constitution imposes institutional limits on the directives of the government and a good constitution will do so in a way that is reflective of the demands of justice.

For example, in 1996 the American Congress passed the Defense of Marriage Act (DOMA) and President Clinton signed it into law. DOMA forced the federal government to not recognize the marriages of homosexual couples whose marriages were recognized in their individual states. It essentially rendered homosexual marriage nonexistent at the federal level. In 2013, DOMA was struck down by the Supreme Court in *United States v Windsor*: “DOMA is unconstitutional as a deprivation of the equal liberty of persons that is protected by the Fifth Amendment.” The Fifth Amendment includes a due-process clause that protects citizens from arbitrary discrimination.

While the court strikes down laws for procedural reasons, i.e. the fact that a law violates the constitution is sufficient reason to strike it down, the constitution itself is an explicitly moral document. In the modern language of liberalism, citizens have the right to due process because they are moral agents, deserving of equal treatment before the law. If the constitution didn't have such a clause, or if the government didn't effectively enforce it, then the directives of the government would not be expert in attaining justice. It doesn't matter whether individuals within the government have expertise about due process (although certainly some do), it matters that the institutional constraints produce directives that accord with due process, a basic demand of justice.

These reflections show the core of truth in the demand that citizens “trust the system”. When a president gets elected that you voted against and begins to implement laws that you think are wrongheaded, you don’t get to break the law just because you disagree. Even if you think the president is mistaken, a trustworthy system will constrain the effects of her mistakes. Egregious enough mistakes will be handled by the courts. Of course, if the courts’ decisions also break from justice then that is a different situation. The point is that it is the institutional constraints, like the constitution, that ground a large part of our trust in the expertise and goodwill of government directives. Most citizens in modern democracies like America do not have more expertise about justice than the constitution itself demonstrates and imposes on government directives in a well-functioning system.

2. Particularity

We can see how laws might apply beyond non-experts and so broadly enough to be sufficient for the achievement of justice by answering another objection. Simmons’ second objection to Rawls’ use of the natural duty of justice is that Rawls is unable to capture the particularity of the duty to obey the law. Again, this is important for Simmons in large part because of his focus on the perspective of the citizen and political obligation. Simmons rightly notes that we often think of political obligation, patriotism, civic-mindedness, and other civic virtues as a package, and he wants to capture that intuition.

The particularity objection to the natural duty of justice comes down to this: if you have a duty to uphold, support, and create just institutions, this should apply as much to your own institutions as any just institution in the world. If this is the case, though, you

don't have a particular duty to obey the laws of your own state, you have a duty to support all the just institutions in the world. One way to do this would be to support your own just institutions but that is not the only way, and likely not the best way. In particular, it often seems that one could achieve justice better by sending one's money to poverty-stricken third world countries and contributing to their basic infrastructure than by paying yet another tax to the governments of the richest countries in the world. This seems both dissonant with our intuitions about having a strong moral bond to our own government and to undermine the cause of justice. After all, if people conform to their reasons by contributing to foreign institutions rather than their own, so many would break the law that the domestic government could not achieve its aims.

My account will of course not justify particular duties to obey the laws of your country because my account does not justify duties to obey the law at all. As far as that goes, I can't answer Simmons challenge. But my account can explain why individuals are often subject to the authority of their government in more domains, which is a form of particularity. I admit it's not a very robust form of particularity, but I think the robust form of particularity Simmons demands is unreasonable. I'll explain these misgivings before turning to the argument in favor of some weak particularity, in the process highlighting some further features of my account.

The main problem I have with Simmons' particularity requirement is its focus on the current world. As Simmons admits, we can easily imagine a world that is composed of nearly or perfectly just societies, and under those conditions it is clear why one has particularly strong moral reasons to obey the law of one's own society. In such an arrangement justice is best achieved by everyone doing their own part in their society, as

the hypothetical assumes all the rest of the world is also doing their own part in their society. But the actual world is not like this; many of the societies of the world are ruled in manifestly unjust ways, which opens up the possibility of better serving justice by helping foreign institutions. Simmons complains,

The natural duty to promote justice would seem then... to require not morally myopic continuing support exclusively for our domestic institutions, but rather that our low-cost assistance in the promotion of justice be carefully budgeted and targeted so as to be most effective in promoting justice.¹⁶

The real world is unjust enough that it seems clear we could better achieve justice by supporting justice in areas of the world outside Western democracies. This is true, but it is no objection.

As I have mentioned several times, I am a cosmopolitan. Simmons' characterization of our current practices as "morally myopic" is exactly right to my mind. A similar criticism to Simmons' particularity objection is sometimes raised against consequentialism. Under current conditions, as Peter Singer has famously argued, consequentialism requires donating massive percentages of our incomes to help the needy in the rest of the world.¹⁷ This is a very demanding moral requirement, so demanding that some people reject it on that ground alone. More worryingly, consequentialism in these circumstances is so demanding that it implies that we often have to sacrifice important elements of our lives for the betterment of others, perhaps including friendships and other personal relationships. But as Elizabeth Ashford has said, the current world is an "emergency situation," and any morality worthy of the title will require helping those in grave need

¹⁶ Wellman and Simmons, p. 165.

¹⁷ Peter Singer, "Famine, Affluence and Morality", *Philosophy & Public Affairs* 1 (1979): pp. 229-243.

when the cost to ourselves is minimal.¹⁸ It would be problematic if consequentialism could not capture some of our intuitions about the demandingness of morality in any world. But the fact that it can't capture our intuitions in this world, where more than twenty thousand people (mostly children) die every day due to preventable, poverty-related causes, is not a problem.

Similarly, it would be problematic if an account of political authority could not explain how, in some ideal circumstances, citizens would have more reason to support their own institutions than others. But the fact that an account of political authority cannot capture particularity in this world, where regimes like North Korea and Syria brutally oppress their own populations, is not a problem. In these conditions, only supporting your own government, or even supporting your own government for the most part, is like watering your lush lawn while people right across the fence die of thirst. At the very least it is morally myopic; plausibly, it is massively unjust.

With my reservations duly noted, I now turn to some considerations that show that an account of evidential modern political authority can capture some of the intuitions behind Simmons' demand. First, once we move away from the idea of justifying a particular duty of obedience, the task becomes much easier. For some of the factors that theorists have appealed to in order to ground particularity fail to support a particular duty but are still good moral reasons for citizens to be particularly concerned with their own laws. If the intuition we are trying to capture is simply that we have more morally-relevant ties to our own political community and thus have some moral reason to prefer it over others, moral reasons that don't ground duties will still be important.

¹⁸ Elizabeth Ashford, "Utilitarianism, Integrity, and Partiality," *The Journal of Philosophy* 97 (2000): pp. 421-439.

For example, in *Crito* the Laws appeal to Socrates' upbringing and life in Athens: Socrates should accept his punishment for breaking Athens' laws because he owes Athens gratitude for the great goods of life it has enabled Socrates to achieve. The idea that gratitude could ground a duty to obey the law has been picked up in modern times as well, but Simmons argues that an appeal to gratitude fails. Gratitude, as important as it might be, can't ground a duty to obey the law. This is right, but my point is that gratitude can still give subjects' good moral reason to be concerned more with the laws of their own government than with others'. This is a weak form of particularity, sure, but it goes a good ways towards explaining the intuition that citizens are more strongly bound to their own governments, the intuition that motivates the particularity objection.

Second, I break with Rawls by maintaining that states have a natural duty of justice as much as individuals, and part of that is creating, supporting, and upholding just foreign political institutions. In this sense my appeal to the natural duty of justice is more in line with Kant's original, quite cosmopolitan thinking than with Rawls. Ideally, much of the work that individuals would be required to do to support foreign institutions would actually be handled by their domestic government. If the government is justified because it is necessary to fulfill our duty of justice and our duty of justice requires supporting just institutions wherever they are found, then the government can only be justified if it helps us support just institutions abroad. And it is worth noting that actual modern governments do this. America isn't a particularly great example, but the Scandinavian governments are better. Part of their legitimacy as states is how they relate to external agents, and that includes to what extent they meet their duty of justice in creating, supporting, and upholding just institutions beyond their own.

I believe it is true that modern governments must use a much higher percentage of their resources towards securing justice for foreigners, and indeed that this is the best way for a large percentage of citizens of those states to discharge their duty of justice. Realism in international relations is simply, obviously, demonstrably false.¹⁹ A good modern government has expertise about how to best use its resources towards the end of global justice, although of course most modern governments aren't good in this sense. America's foreign policy is characterized much more by self-interest than by any real concern for justice, even when those policies are "justified" in the language of assisting others to achieve justice. However, our concern is not whether we can justify the status quo but whether it makes sense that in some circumstance, likely more ideal than our current world, individuals can discharge their duty of justice both domestically and internationally by following the directives of their own government. And this makes sense if governments are reliably assisting the cause of justice internationally.

That said, even in the world as it is today there are two factors that make citizens more particularly bound to their own governments than to foreign governments, in terms of vulnerability to the law. These two factors are characteristics of justice: justice is *contextual* and *indeterminate*. For the rest of this section I'll focus on the contextual nature of justice and push the discussion of indeterminacy until the next section and my treatment of democracy. Considering these factors answers one of the ways Simmons frames the particularity objection. If citizens' duty to achieve justice does not entail they have a particularly important relationship to their own states' laws, it seems to follow that citizens

¹⁹ For the classic rebuttal, see Charles Beitz, *Political Theory and International Relations* (Princeton: Princeton UP, 1979).

are subject to whatever reasonably just system claims to bind them. But this seems clearly mistaken, as Simmons writes:

For nobody, I assume, would seriously argue that if American lawmakers did in fact specify or intend that French citizens (in France) should pay American taxes, abide by American federal law, and so on, that the French would thereby have the same moral duty to obey as do resident American citizens.²⁰

So we need some reason to expect that French citizens are not bound by American laws, even if those laws are reasonably just and claim to bind French citizens.

First, then, justice is contextual. The fact that justice is contextual not only helps explain particularity but shows that the authority of governments will often bind even experts about justice in their daily lives. If this is true, it's starting to look like some modern governments may actually have the authority to bind almost all their citizens with almost all laws.

Justice is contextual in that what is required to meet the demands of justice in one context may not be the same as what is required in other contexts. For example, one of the primary aims of political institutions is to secure the personal safety of subjects. All governments must do so, but how each government should do it depends on the context. What is required to keep citizens safe in the Arctic is much different than what is required to keep them safe in the Amazon. Even if I am an expert about justice in the abstract, such that I know justice requires securing personal safety and everything else, I need to know a lot of empirical facts about the government's actual context of action in order to know what the demands of justice require in this case. But it is plausible that the sort of person who has devoted herself to studying theories of justice is not the sort of person who has studied all the relevant empirical data. So although the government may be less expert about the

²⁰ Wellman and Simmons, *Is There a Duty to Obey the Law?*, p. 163.

demands of justice in general, it is more expert about the demands of justice in its own context. The government has the relevant expertise to have authority over me.

If this is true of justice experts, then it is going to be true of almost all of the population. Even those people that have expertise in individual, justice-relevant domains will be bound by all the laws in domains that they lack expertise in, which will be most domains. This is why we should think it plausible that some modern governments could have authority of the sort that binds a level of the population sufficient to establish a fair and effective system of punishment, so ultimately to achieve justice.

The reason French citizens in France wouldn't be bound by American laws that claimed to bind them is that American laws don't have authority over them. The lack of authority is due to the American government's lack of contextual expertise: the American government doesn't know enough about the particular physical, historical, and cultural French context that shapes the demands of justice in France. Thus, directives that were intended to bind French citizens in France would often fail to help French citizens conform better to their duty of justice, especially as compared with the possibility of following the directives of the reasonably just French government. The American government does not meet the expertise condition on genuine evidential authority with respect to French citizens, so those citizens aren't bound by American law and can't be made vulnerable for disobeying American law. In the next section I'll also argue that the American government doesn't meet the trustworthiness condition on genuine authority with respect to French citizens.

In the absence of a reasonably just French state, it's more plausible that a reasonably just foreign government could have justice expertise with respect to individuals

in France. This still doesn't show that French individuals would be vulnerable to punishment for disobeying American law; for one thing, the lack of contextual expertise would mean that only some of the most narrow and basic demands of justice could reasonably be imposed on foreigners. Further, if the achievement of justice requires a fair system of punishment, then the fact that the American government has no punishment institutions in France would mean that many people would disobey American directives, such that those directives do not help even those who obey to achieve justice. For example, if the American government directs French citizens to drive on the left, most will ignore that and continue driving on the right because there's no effective and fair American system of punishment in France. In those conditions, even those who were disposed to obey the American directive shouldn't; the fact that everyone else is driving on the right gives them decisive moral reason to also drive on the right. So following the American directives doesn't help French citizens better conform to their preexisting reasons. Perhaps more importantly, the American government would still suffer from a severe deficit of trustworthiness, as I explain below.

The fact that governments have contextual expertise shows that their wide-ranging authority can extend even to experts about justice in the abstract, but it doesn't show that the standing of justice experts and non-experts with respect to the law is identical. Here it is important that, following Raz, my approach to authority is piecemeal. The grounds of authority are whether the commands of the authority will help a particular individual conform better to her preexisting reasons in a particular domain. We can thus evaluate the authority of governments with respect to all individuals and all domains; you may be bound in some domains but not others, while I am bound by different domains than

you, and a third person is not bound at all. Actual governments often make mistakes that undermine their authority in particular domains, especially with respect to justice experts.

Consider the issue of LGBTQ marriage. This is a basic demand of justice, a demand that governments must meet regardless of their context.²¹ Not to allow LGBTQ individuals to have legally recognized marriages constitutes arbitrary legal discrimination because sexual orientation is a morally arbitrary characteristic. It might seem that, due to governments' contextual justice expertise, I'm claiming that those discriminatory laws bind all citizens, even those who recognize the arbitrariness. Because of my piecemeal understanding of authority, however, that's not right. There are some mistakes of justice that the government can make that would not undermine their authority precisely because they have contextual justice expertise with respect to many citizens. But discrimination on arbitrary grounds is not one of those mistakes because it is a basic demand of justice. If authority were all-or-nothing, then mistakes about basic demands of justice might mean the government lacks authority over justice experts in all domains. On the piecemeal understanding, though, we can claim that the government lacks authority specifically with respect to the domain of LGBTQ treatment.

For example, if I am a member of the clergy in a state where such marriages are outlawed and I am also an expert on the demands of justice, I am not objectively morally vulnerable to punishment if I marry an LGBTQ couple, even if I manage to trick the government into legally recognizing the marriage. If I am punished for this, I have a

²¹ Sometimes under conditions of extreme resource scarcity, governments must make choices between which aspects of justice to meet. In those sorts of contexts it's plausible to me that some deviations from the ideal standards of justice don't undermine authority. For this discussion, though, I'm referring primarily to differences in geography and culture, rather than resources, as modern states of the sort we're considering are resource rich.

justified moral complaint against the government. Now, as far as the real world goes, the government doesn't recognize a permission for me to avoid that punishment, so if I am caught I will likely actually be punished. But because of the basic injustice and resultant lack of authority, whatever reasons I have not to break this law are merely prudential. In other domains of my life, though, I still have good reason to believe that the government's directives are more reliable than mine and the government has authority over me. It only lacks authority in this particular domain because of its basic injustice in that domain.

There are even more nuances in determining whom and to what extent actual laws bind. I'm going to discuss one more, but it's important to note that on my theory of authority, evaluating whether a particular agent is bound by a particular law by a particular government is never going to be easy. This is especially true because of the nuance I want to touch on, namely the effect of injustice on the trustworthiness condition.

A basic injustice such as arbitrary discrimination has deleterious effects on a government's trustworthiness, not just on their expertise. If the government can discriminate against one group arbitrarily, there's nothing to stop it from discriminating against other arbitrarily chosen groups. If that's the case, then the government might be so untrustworthy that it loses all authority. Now, it would be too quick to claim that any arbitrary treatment renders the entirety of the government's directives so untrustworthy that they don't bind. As in the LGBTQ case, it is sometimes obvious that the government is arbitrary about a specific characteristic for specific reasons that don't translate to other traits. On the other hand, if a government discriminates against women and specific religions and homosexuals and political opponents, it shows that they misunderstand the basic imperative of government to treat subjects equally, according to morally non-

arbitrary characteristics. That sort of government is so untrustworthy that it plausibly lacks all authority.

The government can also lose authority over specific groups of citizens by discriminating against them across many domains, such that members of that group are not vulnerable to punishment for disobeying the law. A clear example of this in the United States is the black underclass. Many members of the underclass have strong evidence of the lack of goodwill of both the federal and their more local governments, evidence that is reinforced on a daily basis. They see intimidation and indifference from government representatives, friends and family killed by endemic violence the government fails to address, systematic discrimination in the criminal justice system, denial of economic and educational opportunities, and so forth.²² All of this is good evidence that the government does not have their interests in mind when making laws; the government lacks goodwill, a central component of trustworthiness, so can't have authority over them. Especially concerning is the systematic injustice of the criminal justice system, wherein blacks are much more likely to be arrested, convicted, and punished more harshly for the same crimes as whites.²³

The fact that the government lacks authority over many members of the black underclass because it fails to meet the trustworthiness condition does not mean those members are never morally vulnerable to punishment. It's not *carte blanche*. Instead, a plausible consequence of this claim may be that they are still vulnerable to punishment for

²² For a recent philosophical treatment of this problem, see Anderson, *The Imperative of Integration*.

²³ For example, see Michelle Alexander, *The New Jim Crow: Mass Incarceration in the Age of Colorblindness* (New York, NY: The New Press, 2010).

violating *mala in se* laws but not for breaking (many) *mala prohibita* laws.²⁴ However, due to the systemic racial injustice of American criminal justice institutions, it may also be the case that the actual American government wrongs those individuals when it punishes them despite the fact that they committed punishable violations.

My point is that failures of expertise and trustworthiness by the government and their relationship to genuine authority are complicated matters, particularly on a piecemeal approach. Sometimes injustices can be traced quite particularly and so only undermine the government's authority in specific, narrow domains. But sometimes injustices are so wide-ranging that they infect other domains and nullify the government's authority more generally.

3. Democracy

At this point I have made the case that Western democracies plausibly meet the expertise condition, due to their institutional nature and contextual expertise, and the deference condition, due to the fact that they uniquely assist in meeting the important duty of justice. The acceptance condition is easily met, requiring only that the agent intends to bind practically as the law clearly does, and the precedence condition is met because the governments' contextual justice expertise is precedent for citizens with respect to meeting their duty of justice. The only remaining condition is trustworthiness. In this section I address trustworthiness, and bolster the case for the expertise of modern governments, by

²⁴ Some *mala prohibita* laws change empirical conditions, due to widespread compliance, such that vulnerability to punishment is also appropriate for breaking those laws. The law against driving on the left in America comes to mind. Thanks to Ian MacMullen for pushing me on this.

addressing what may have seemed to you an obvious hole in my account, namely the place of democracy.

A central tenet of modern commonsense political philosophy is that democracy is a necessary condition on political authority. The history of political organization in the modern consciousness is the move away from anti-democratic, tyrannical regimes to governments “of the people, for the people, by the people”. Democracy is claimed to be the key to international peace, the way for peoples to overthrow oppressors and join the global community. The argument in favor of democracy is a strong one. Buchanan, for example, argues that,

[I]f we take the equality of persons seriously, then a political order that not only honors the commitment to equal regard by respecting all citizens’ human rights, but also does so by political processes that themselves express this commitment to equality by being democratic, would seem to provide the best answer available to the problem of reconciling political power and equality.²⁵

Taking the moral equality of persons seriously is a necessary feature of any plausible morality; Buchanan calls equality of persons “the most fundamental moral principle of all”.²⁶ An account of political authority that rejects democracy disrespects people as free and equal members of the moral community. Expertise and democracy, however, seem diametrically opposed. We don’t appeal to democracy because it makes better decisions, and when we appeal to experts it seems to be precisely because uninformed opinions are insufficient.

Four qualifications will help us see the force of this concern for my account. First, recall again that we are concerned with expertise as an emergent property of government

²⁵ Allen Buchanan, *Justice, Legitimacy, and Self-Determination* (Oxford: Oxford UP, 2004): p. 250.

²⁶ *Ibid.*, p. 253.

directives due in large part to institutional constraints on individuals in institutional roles. If my account claimed that governments have authority only when they are under the control of experts, it would be decidedly undemocratic. But that's not what my account claims, so there is the possibility that democracy itself is one of the institutional constraints that leads to government directives having emergent expertise.

Second, I am going to ignore the intrinsic value of democracy and focus on its instrumental value, particularly with respect to achieving justice and having genuine evidential authority. It may well be the case that respecting individuals requires that the government employ democratic decision-making procedures. I think this is true but also that it is an issue of political legitimacy rather than political authority. It makes sense to ignore the intrinsic value of democracy in this context because while democracy may be a necessary condition on political legitimacy, nobody plausibly claims that it is a sufficient condition. At its core democracy is a procedural thesis: it tells us that we should use a democratic method to make political decisions, but it does not tell us what those decisions should be.²⁷ If the procedure results in patently and egregiously unjust decisions, they cannot bind regardless of the fact that they were democratically produced. The majority cannot oppress minorities and then claim the oppression was justified because it was done with laws that the majority made. So while it may be the case that democracy is instrumental towards achieving justice, it's not the case that instituting democratic procedures simply entails that justice is achieved.

²⁷ I am not denying that democracy includes much more than a procedural thesis, as emphasized, e.g., by Anderson, *The Imperative of Integration*, ch.5. But the procedural thesis is a central component; I take it that a robustly democratic culture and civil society wouldn't serve the valuable functions they do without the actual procedures enabling the people to express their values in the operation of their government.

Third, it is perfectly consistent with my theory to claim that the intrinsic value of democratic procedures sometimes entails that less just democratic decisions should be preferred to more just non-democratic decisions. This is due to the deference condition, which if you recall claims that A cannot have authority over B if B has decisive reason to make the decision on her own judgment. The example I gave to illustrate this was the responsabilization of children: even though parents may be able to choose more nutritious foods, at some point allowing children to make their own decisions (for example what to eat for lunch) is more important than the result of those decisions. It is the process of making decisions for oneself that matters here. The parent, even though she could make a better decision, should not interfere. Failure to respect these sorts of reasons leads to helicopter parenting and overprotected children who cannot act well in the world on their own.²⁸ Even in circumstances where parents have expertise relative to their child, the deference condition would not be met.

The same argument can be made in favor of groups of people making their own decisions about some issues, even when those decisions might be worse (although this ignores the point that the members of the government are also part of the group). Just like there are reasons for people to decide for themselves as individuals, there are reasons for people to decide for themselves as collectives. As Buchanan noted, these are weighty reasons because they originate in the fundamental moral equality of all people as well as their rights and interests in joining together into larger communities. The reasons favor a particular decision-making procedure, letting the people decide collectively, over others, like forcing people to defer to an authority. On the plausible assumption that democracy is

²⁸ See Buster Bluth.

quite important, the deference condition will not be fulfilled over a broad range of questions and decisions should be left up to the people rather than to experts.²⁹ Even though expertise grounds authority, we often have good reason to follow democratic procedures, as when there is especially important reason for the public to decide for itself.

Fourth, expertise is only relevant where the relevant duties are determinate. Consider the issue of driving on one side of the road. There is a determinate obligation for everyone to drive on the same side, without which driving would be incredibly risky, but there is not a determinate side that everyone must drive on. Governments can choose either the right or the left, but they must choose one over the other. Expertise is relevant for determining that some side has to be chosen but is silent with respect to which side should be chosen. Moral demands are often indeterminate in this way.

In fact, we might expect much of the political domain to be indeterminate. This is perhaps not obvious if which side of the road to drive on is our paradigmatic case. A better case is property. Here is Anna Stilz clearly and concisely arguing property is morally indeterminate:

The basic thought here is that while a principle of equal freedom provides us some information about what just property distributions look like, the principle's content is underspecified, and therefore cannot be directly applied. The equal freedom principle suggests that whatever system of property we implement, it ought to be consistent with everyone's possession of a zone of freedom that is guaranteed against others' coercive interference. Nevertheless, many possible systems of property—collective allocation, market socialism, unfettered private ownership—are potentially consistent with that sense of equal freedom. And under each one of these many possible systems, there will again be many possible particular rules consistent with everyone's freedom—rules about the precise bundle of claims conferred by ownership, about how exchange is to be regulated, about which objects belong to which particular persons. And finally, any system of

²⁹ Note that "broad" is a relative term here. An appeal to democracy is not a strong enough reason to allow the violation of human rights, for example.

property will also have to include some aspects that are wholly conventional: rules about what precise formalities are required to conclude a contract, exactly how long a statute of limitations to institute, down, indeed, to what side of the road to drive on.³⁰

So moral reasons, for example given by a principle of equal freedom, determine that some systems of property are ruled out, for example complete state ownership. But they do not uniquely endorse one system over the others, let alone the details of one version of one system over other versions.

Expertise is not required to answer these questions, although expertise is required to make sure that the answers fall within determinate bounds. But when choosing among permissible options to give specific content to an indeterminate moral demand, we can't appeal to better and worse options. Instead, it makes sense to look to procedural reasons; we should choose the democratically authorized option. So we have good reason to suspect that democratic procedures should be instituted in any political institution, because they are the best procedure for choosing among options that could equally well meet an indeterminate demand of justice, like what property scheme to use.

With these qualifications in mind, I now turn to the case for democracy's immense, perhaps unique, instrumental contributions to both trustworthiness and expertise. In the end, it is plausible that no modern government can possess authority without being democratic, thus affirming the strong pro-democracy intuitions in modern commonsense political thinking.

Recall that trustworthiness is composed of two elements, competence and goodwill. On my account of authority, competence is ensured if the government meets the expertise condition, so I'll address that component of trustworthiness below. Here I will focus on

³⁰ Anna Stilz, *Liberal Loyalty* (Princeton: Princeton UP, 2009): p. 40.

democracy's contribution to goodwill. In this context goodwill means willingness to take account of and protect the interests of the subject. Democratic procedures help protect citizens' interests by giving the citizens a voice in assessing whether their interests are being sufficiently protected and promoted. If a president keeps acting against the (perceived) interests of the citizenry, they will hold her accountable by electing someone else, someone they believe is more trustworthy. Of course they may be mistaken, but non-democratic governments have a much worse record of accountability to citizens' interests.

A good proxy for goodwill is the republican concern with non-domination. A dominates B just in case A has the capacity to arbitrarily interfere with B's interests, where arbitrariness is determined by the degree of influence or control B's interests have on A's decision-making.³¹ Democratic procedures force the interests of the people into the decision-making of governments, to greater and lesser degrees. Governments must be trustworthy, which requires goodwill, which requires non-arbitrariness in accounting for the interests of each individual.

In previous chapters I said nothing about what actual characteristics a government must have to achieve the level of trustworthiness necessary to ground its authority. I only noted that it is *very* difficult, given how vulnerable we are to political agents. I then gave some arguments suggesting that, quite schematically, political agents are able to take advantage of their institutional nature to achieve a level of competence and goodwill that would be otherwise impossible. Here I can add a substantive claim to this argument. It is plausible to me that without being democratic, no modern government could establish sufficient goodwill to be sufficiently trustworthy to have the kind of sweeping authority

³¹ Philip Pettit, *Republicanism* (Oxford: Oxford UP, 1997).

governments claim. We are steeped in stories of the abuses of power, catastrophic history lessons in government run amok. There is a reason commonsense modern political philosophy sees democracy as the uniquely legitimate form of modern political organization.

This suggests another answer to Simmons' objection regarding American and French law. Above I suggested that French citizens wouldn't be bound because the American government lacks the appropriate contextual expertise. Because the American government is also not democratic with respect to the French citizens, there will be a lack of trustworthiness. Without institutional constraints like regular free elections, French citizens have no reason believe that the American directives are being guided by goodwill: American directives aren't being constrained by French interests and so dominate the French.

Although the institutional nature of governments can be exploited in favor of trustworthiness, it also gives us reasons to doubt their goodwill and competence. Institutions encourage some phenomena that undermine trustworthiness, for example the diffusion of responsibility that occurs when people feel their part in a collective action is so minor, or so controlled by other forces, that they are not morally responsible for their actions. Feeling this loss of responsibility, they are more likely to condone and participate in collective action regardless of whether that action is good, in the political case meaning whether it really serves the public interest. One very important way that political institutions can be trustworthy is with democratic procedures that act as a counterweight to these sorts of institutional problems. Democratic procedures are the best check on the

political agent acting to secure ends other than the interests of its citizens, i.e. acting arbitrarily.

So democracy is immensely, perhaps uniquely, valuable in its ability to secure non-arbitrariness and so demonstrate to citizens the goodwill and trustworthiness of the government by institutionally accounting for and protecting each individual's interests. In a similar vein, although not as strongly, democracy contributes to expertise. As advocates of deliberative democracy and others have emphasized, democratic procedures can often secure better outcomes. Institutions that incorporate democratic procedures will sometimes, perhaps even often, be more competent.

The famous case is the UK Cumbrian sheep case.³² After the nuclear disaster at Chernobyl, parts of the UK began seeing radioactive rainfall, which effected the grass and ultimately the sheep that fed on the grass. Scientists from the federal government began imposing restrictions on the local sheep farmers in order to contain the radioactive consequences. The scientists assumed that their expertise was global and could be applied equally well to the Cumbrian sheep farmers as to their standard cases. Their policies ultimately led to worse outcomes precisely because they ignored the local sheep farmers and imposed abstract standards. In my terminology from above, the sheep farmers had contextual expertise relevant for the application of general standards that the scientists did not have. Had the government's procedures been more democratically informed, the results would have been better.

Some people claim that democracy is completely justified on epistemic grounds, but I don't need a claim nearly that strong. The point is simply that democratic procedures

³² Brian Wynne, "Sheepfarming After Chernobyl: A Case Study in Communicating Scientific Information", *Environment* 31 (1989): pp. 10-15.

sometimes lead to more competent and so more reliable judgments, contributing to the government's expertise and ability to meet the modified normal justification thesis. The most common argument in support of the claim that democracies make governments more competent relies on the Condorcet Jury Theorem. But there are limitations to applying this theorem, and good reasons to doubt it applies in modern democracies.³³

A better argument from Elizabeth Anderson relies on the works of John Dewey. She brings out three epistemic aspects of democracy that are well explained by Dewey's account.³⁴ These are diversity, discussion, and dynamism. I don't have the space to fully explain the argument, but two examples can illustrate the basic case. The diversity claim notes that democracy (assuming universal suffrage) brings together diverse groups, cutting across class, geographic, racial, gender, ethnic, cultural, and other lines. These groups are also epistemically diverse, and epistemic diversity leads to better outcomes. The sheep farmers knew more about their particular context than the government's experts, and had their knowledge been institutionally available, the outcome of the government's policies would have been better.

Hélène Landemore also appeals to the diversity of democratic procedures to make the epistemic case for democracy.³⁵ Landemore, drawing on studies of deliberative decision-making in political science, argues that in liberal conditions democracy radically improves the outcomes of government decisions precisely because of the extensive cognitive diversity of the electorate. She argues, "under the right conditions and all things being equal otherwise, what matters most to the collective intelligence of a problem-

³³ For sample critiques, see Elizabeth Anderson, "The Epistemology of Democracy", *Episteme* 3 (2006): pp. 8-22 and David M. Estlund, *Democratic Authority* (Princeton: Princeton UP, 2008).

³⁴ Anderson, "The Epistemology of Democracy".

³⁵ Hélène Landemore, *Democratic Reason* (Princeton, NJ: Princeton UP, 2013).

solving group is not so much individual ability as the number of people in the group.”³⁶ The argument for this claim is complex and I certainly don’t want to rest my case for genuine authority on it. But if it is true, it makes my point: it’s not about whether individual experts are in the government but about the contextual knowledge of millions of many individuals under good institutional constraints that result in directives with emergent expertise.

Democratic procedures not only add considerable diversity, they are dynamic. Dynamism is Anderson’s alliterative term for feedback and change mechanisms. The commands of political authorities are intended to solve practical problems in the lives of individuals. The success of policies is best evaluated by the people whose lives are affected, and holding decision-makers accountable for failed policies is a primary way institutional mechanisms can help secure better outcomes. Policies that allow for external evaluation and change will be better, and democracy helps secure these outcomes. This claim is about more than voting, though, for other components of a democratic society, like a free press and active dissent, contribute to epistemic competence. Feedback mechanisms significantly contribute to both competence and goodwill.

The point is simply that democracy can increase competence, and so make it more likely that democratic governments will meet the expertise and trustworthiness conditions than non-democratic governments. I leave aside the many interesting questions remaining, such as how much competence is increased, under what conditions democracy also entails epistemic losses, what specific mechanisms increase competence, and so on. These would be more relevant for a theory that claimed democracy was justified purely on epistemic grounds, but that is not my claim. There is simply good reason to believe that democracy

³⁶ Ibid., p. 104.

plays a significant and perhaps unique role in modern governments' meeting the conditions for genuine political authority.³⁷

If this is right, then democratic authorization procedures of some sort will be a necessary feature of any genuine modern political authority. Intuitively this is the correct result, especially given our exclusive focus on the modern political context. Democracy is a plausible condition on modern political authority because it is the primary way to establish accountability and so trustworthiness with respect to goodwill. Further, access to the diverse knowledge of huge populations aids the governments' judgments by increasing their contextual expertise and ultimately helping citizens better conform to their preexisting reasons. Thus, as Buchanan argues, where we have the resources and capability to use democratic procedures we should.³⁸ It is not a necessary condition on genuine political authority of all types, but given our modern history and resources, it will be required of most, or perhaps all, modern political authorities. On the whole, my account leaves ample space for democracy and for the expression of democratic values.

Conclusion

This completes the case for evidential governmentalism. All people have a natural duty of justice to create, support, and uphold just institutions that meet Rawls' two principles of justice. This duty is extremely important, important enough that people can be permissibly punished for not fulfilling it. One of the necessary, constitutive components of these

³⁷ There is a nuanced, and perhaps more powerful, Rousseauian argument to the conclusion that democracies are necessary for the formation a collective will, and it is only with a collective will that we can define what is in the public interest and so which problems ought to be addressed by political authority. Democracy would thus be a necessary condition on competence when addressing collective problems. Unfortunately I don't have the space to address that argument here.

³⁸ Buchanan, *Justice, Legitimacy, and Self-Determination*, pp. 249ff.

institutions is a fair system of punishment, which requires that the government have the liberty to punish subjects and correlatively that subjects are vulnerable to punishment. In order to make subjects vulnerable to punishment, the government's commands must change subjects' subjective duties with respect to how they meet the duty of justice. The power to issue commands that changes' others subjective duties of the relevant sort is grounded primarily in expertise.

Some modern governments, particularly Western democracies, have expertise about how to meet the duty of justice, as we see in how close their institutions and outcomes get to meeting Rawls' two principles. Their judgments have the emergent property of justice expertise due to institutional constraints, of which the constitution is one particularly important and prominent example. They especially have expertise relevant to their average citizens, who are often benighted about the true demands of justice. Further, because they have the sort of contextual expertise required to achieve justice in particular circumstances, they have justice expertise over most of their population, benighted or otherwise. Democratic institutions considerably bolster this contextual expertise. Modern governments also must be sufficiently trustworthy to justify the level of deference they require in the very important domains they govern. Again, democratic procedures and other institutional mechanisms are the primary reason governments can achieve the required level of trustworthiness.

Under these conditions, most of the commands of some modern governments will bind most of their subjects in most domains, constituting preemptive, content-independent evidential reasons that make subjects objectively morally vulnerable to punishment for disobeying. This allows the government to institute a fair system of punishment that,

because vulnerability is widespread enough, can be sufficiently effective for the government to coordinate behavior and achieve justice. Some governments have genuine evidential modern political authority over most of their subjects in most domains.

Evidential governmentalism is true. Philosophical anarchism is false.

Works Cited

- Alexander, Larry. "Law and Exclusionary Reasons." *Philosophical Topics* 18 (1990): 5-22.
- Alexander, Larry, Kim Kessler Ferzan, and Stephen J. Morse. *Crime and Culpability: A Theory of Criminal Law*. Cambridge: Cambridge UP, 2009.
- Alexander, Michelle. *The New Jim Crow: Mass Incarceration in the Age of Colorblindness*. New York, NY: The New Press, 2010.
- Anderson, Elizabeth. "The Epistemology of Democracy." *Episteme* 3 (2006): 8-22.
- . *The Imperative of Integration*. Princeton: Princeton UP, 2010.
- Anscombe, G.E.M. "On the Source of the Authority of the State." Raz.
- Applbaum, Arthur Isak. "Legitimacy without the Duty to Obey." *Philosophy and Public Affairs* 38 (2010): 215-239.
- Ashford, Elizabeth. "Utilitarianism, Integrity, and Partiality." *The Journal of Philosophy* 97 (2000): 421-439.
- Austin, J. L. *How to Do Things with Words*. Oxford: Oxford UP, 1955.
- Baier, Annette. "Trust and Antitrust." *Ethics* 96 (1986): 231-60.
- Beitz, Charles. *Political Theory and International Relations*. Princeton: Princeton UP, 1979.
- Bratman, Michael. "Practical Reasoning and Acceptance in a Context." *Mind* 101 (1992): 1-16.
- Buchanan, Allen. *Justice, Legitimacy, and Self-Determination*. Oxford: Oxford UP, 2004.
- Caplan, Arthur. "Moral Experts and Moral Expertise: Does Either Exist?" *If I Were a Rich Man Could I Buy a Pancreas?* Bloomington: Indiana UP, 1992.
- Copp, David. "The Idea of Political Legitimacy." *Philosophy and Public Affairs* 28 (1999): 3-45.
- Darwall, Stephen. "Authority and Reasons: Exclusionary and Second-Personal." *Ethics* 120 (2010): 257-278.
- Ehrenberg, Kenneth. "Critical Reception of Raz's Theory of Authority." *Philosophy Compass* 6 (2011): 777-785.
- . "Joseph Raz's Theory of Authority." *Philosophy Compass* 6 (2011): 884-94.
- Elga, Adam. "Reflection and Disagreement", *Nous* 41 (2007): 478-502.
- Essert, Christopher. "A Dilemma for Protected Reasons." *Law and Philosophy* 31 (2012): 49-75.
- Estlund, David M. *Democratic Authority*. Princeton: Princeton UP, 2008.
- Friedman, Richard B. "On the Concept of Authority in Political Philosophy." Raz.
- Goldman, Alvin I. "Experts, Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63 (2001): 85-110.
- Hardwig, John. "Epistemic Dependence." *The Journal of Philosophy* 82 (1985): 335-49.
- . "The Role of Trust in Knowledge." *The Journal of Philosophy* 88 (1991): 693-708.
- Hart, H. L. A. "Commands and Authoritative Reasons." *Essays on Bentham*. Oxford: Clarendon Press, 1982.
- . *The Concept of Law*. Oxford: Oxford UP, 1961.
- Himma, Kenneth. "Just 'Cause You're Smarter than Me Doesn't Give You a Right to Tell Me What to Do: Legitimate Authority and the Normal Justification Thesis." *Oxford Journal of Legal Studies* 27 (2007): 121-150.
- Hobbes, Thomas. *Leviathan*. Ed. Richard Tuck. Cambridge: Cambridge UP, 1991.

- Hohfeld, Wesley Newcomb. *Fundamental Legal Conceptions as Applied to Judicial Reasoning*. New Haven: Yale UP, 1919.
- Huemer, Michael. *The Problem of Political Authority: An Examination of the Right to Coerce and the Duty to Obey*. New York: Palgrave Macmillan, 2013.
- Hurd, Heidi M. "Challenging Authority." *The Yale Law Journal* 100 (1991): 1611-77.
- Jackson, Frank. "Decision-theoretic Consequentialism and the Nearest and Dearest Objection." *Ethics* 101 (1991): 461-82.
- Jones, Karen. "Trust as an Affective Attitude." *Ethics* 107 (1996): 4-25.
- Ladenson, Robert. "In Defense of a Hobbesian Conception of Law." *Raz*.
- Landemore, H el ene. *Democratic Reason*. Princeton, NJ: Princeton UP, 2013.
- List, Christian and Philip Petit. *Group Agency: The Possibility, Design, and Status of Corporate Agents*. Oxford: Oxford UP, 2011.
- McMahan, Jeff. *Killing in War*. Oxford: Oxford UP, 2011.
- Mian, Erman. "The Curious Case of Exclusionary Reasons." *Canadian Journal of Law and Jurisprudence* 15 (2002): 99-124.
- Moore, Michael S. "Authority, Law, and Razian Reasons." *Southern California Law Review* 62 (1988): 827-896.
- Nisbett, Richard E. and Timothy D. Wilson, "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84 (1977): 231-59.
- Parfit, Derek. *On What Matters*. Oxford: Oxford UP, 2011.
- Pettit, Philip. *Republicanism*. Oxford: Oxford UP, 1997.
- Quinn, Warren. "The Right to Threaten and the Right to Punish." *Philosophy and Public Affairs* 14 (1985): 327-373.
- Rawls, John. *A Theory of Justice*, Rev. ed. Cambridge: Harvard UP, 1999.
- . *Political Liberalism*. New York: Columbia UP, 1993.
- . "Two Concepts of Rules." *The Philosophical Review* 64 (1955): 3-32.
- Raz, Joseph, ed. *Authority*. New York: NYU Press, 1990.
- . *Between Authority and Interpretation: On the Theory of Law and Practical Reason*. Oxford: Oxford UP, 2010.
- . "Facing Up: A Reply." *Southern California Law Review* 62 (1988): 1153-1235.
- . *From Normativity to Responsibility*. Oxford: Oxford UP, 2011.
- . *Practical Reasons and Norms*. Princeton: Princeton UP, 1975.
- . *The Authority of Law*. Oxford: Oxford UP, 1979.
- . *The Morality of Freedom*. Oxford: Oxford UP, 1986.
- . "The Problem of Authority: Revisiting the Service Conception." *Minnesota Law Review* 90 (2006): 1003-1044.
- Regan, Donald H. "Authority and Value." *Southern California Law Review* 62 (1989): 995-1096.
- Sidgwick, Henry. *The Methods of Ethics*. London: MacMillan, 1874.
- Simmons, A. John. "The Anarchist Position: A Reply to Klosko and Senor." *Philosophy and Public Affairs* 16 (1987): 269-70.
- Singer, Peter. "Famine, Affluence and Morality." *Philosophy & Public Affairs* 1 (1979): 229-243.
- Stilz, Anna. *Liberal Loyalty*. Princeton: Princeton UP, 2009.
- United Nations. Human Rights Council. *Report of the Commission of Inquiry on Human Rights in the Democratic People's Republic of Korea*. New York: United Nations, 2014.

United States v Windsor. 570 U.S. 12. Supreme Court of the United States. 2013.

van der Vossen, Bas. "On Legitimacy and Authority." *Res Publica* 14 (2008): 299-302.

Wellman, Christopher Heath. "Liberalism, Samaritanism, and Political Legitimacy." *Philosophy and Public Affairs* 25 (1996): 211-237.

Wellman, Christopher Heath and A. John Simmons. *Is There a Duty to Obey the Law?* Cambridge: Cambridge UP, 2005.

Williams, Bernard. "Internal and External Reasons." *Moral Luck*. Cambridge UP, 1980.

Wolff, Robert Paul. *In Defense of Anarchism*. New York: Harper & Row, 1970.

---. "The Conflict Between Authority and Autonomy." *Raz*.

Wynne, Brian. "Sheepfarming After Chernobyl: A Case Study in Communicating Scientific Information." *Environment* 31 (1989): 10-15.

Appendix: The Conceptual Coherence of Exclusionary Reasons

Many theorists have serious doubts about exclusionary reasons: “exclusionary force is conceptually problematic. It is difficult to understand, it is difficult to maintain, and it is difficult to see the point of it.”³⁹ In this appendix I address the concern that the very idea of exclusionary reasons is conceptually incoherent, considering arguments from Christopher Essert and Heidi Hurd.⁴⁰

Essert’s concern focuses on the reasons that are excluded. Let’s say an authority commands me to go to the gym. I have preexisting reasons to go to the gym, like the benefits of exercise, my desire to stay on a habitual gym schedule, and so on. I also have preexisting reasons not to go to the gym, like the pain of exercise, the time I could spend doing other activities, and so on. The authority’s command excludes some of these reasons; the question is which.

As noted in the previous section, in earlier writing Raz characterizes the second-order reason as excluding *all* the other reasons from the subject’s practical deliberation. So in the gym case my reasons to not go to the gym, like the pain of exercise, get excluded, but so do my reasons to go to the gym, like the benefits of exercise. Call this the broad interpretation of the scope of exclusion. However, more recently Raz has modified his position to say that only reasons counting in favor of acting in a manner not commanded are excluded. So the pain of exercise is excluded from my deliberation because it supports not going to the gym, but the benefits of exercise are not excluded because they support doing as commanded, i.e. going to the gym. This is the narrow interpretation. Essert claims

³⁹ Emran Mian, “The Curious Case of Exclusionary Reasons”, p. 99.

⁴⁰ Christopher Essert, “A Dilemma for Protected Reasons”, *Law and Philosophy* 31 (2012): pp. 49-75; Hurd, “Challenging Authority”.

that exclusionary reasons are incoherent on either interpretation and so presents a dilemma for exclusionary reasons.

The first horn of the dilemma rests on the broad interpretation. The problem (the sharp point of this horn of the dilemma) is that it seems perfectly acceptable to do as commanded for whatever reason, as I have been stressing. If I decide to go to the gym because I want to get healthier, and ignore the fact that I was commanded to go, I am still doing as commanded. And this seems sufficient to count as obedience to the authority. I do not have to go to the gym simply because I was commanded to, but this is what is required if all my reasons (other than the command-reason) are excluded.

The second horn of the dilemma rests on the narrow interpretation. Essert argues that the problem on this horn is that subjects are forced to double-count reasons. Here is the idea: when reasons are excluded, it is still up to the subject to go through her deliberation and act. Of course, the exclusionary reasons have rigged the game, as it were, by leaving only reasons counting in favor of the commanded act on the table for the subject to consider. In the gym case, the pain of exercise is excluded as a reason for action, but the benefits of exercise remains. What also remains, and this is the problem, is the fact of the command.

Recall that authoritative directives give a preemptive reason with two parts: the exclusionary, second-order reason as well as the first-order reason. This first-order reason, the fact that an authority has commanded the subject to act, is one of the reasons the subject may consider in her practical deliberation. But good authorities conform to the dependence thesis: the authorities' judgments are based on and take account of the

preexisting reasons to act.⁴¹ In the gym case, the fact that exercise will benefit me is one of the preexisting reasons, and the authority must have taken that fact into account in her directive for me to go to the gym. If the subject deliberates about how to act and tries to weigh the reason grounded in the benefits of going to the gym in addition to the reason grounded in the command, she will double-count the benefits of going to the gym. For the benefits have already been counted in the authority's judgment.

This kind of double-counting may not seem pernicious because it is just extra force in favor of the commanded act, and that's all the subject is permitted to do anyway. But this ignores the fact that authorities have domain-specific jurisdiction, and only reasons that fall within her jurisdiction are excluded. In the gym case, perhaps the authority's jurisdiction only extends to my physical health. We can modify the case so that there are reasons outside the authority's jurisdiction that count against going to the gym; perhaps my sister is sick and I want to bring her some soup. In this case, as I am balancing the reasons for and against going to the gym, double-counting is problematic. Let's implausibly stipulate that the reason to visit my sister has a weight of 5, the reason based in the physical benefits of exercise a 2, and the reason based in the command a 4. If I count only the sister and command reasons, then I visit my sister because the reason of weight 5 outweighs the reason of weight 4. But if I double-count by also adding in the exercise reason, I go to the gym because the combined reasons in favor of going to the gym at 6 outweigh the reason to see my sister at 5. This case is, of course, overly stylized and the strengths I assigned may seem intuitively preposterous, but it simply serves to illustrate

⁴¹ Raz, MF, p. 42ff.

that double-counting is problematic primarily because of reasons which fall outside the authority's jurisdiction.

As presented, Essert's dilemma is problematic because each horn depends on a different interpretation of exclusionary reasons. The first only works against the motivational interpretation and the second, as Essert spells it out, assumes the justificatory interpretation. That said, I think the problem in the second horn can be modified so it works against the motivational interpretation as well. This modified dilemma would then present a problem for Raz, but not for me because I reject the motivational interpretation in favor of the justificatory interpretation. On the justificatory interpretation, the first horn is not a worry.

The first horn claimed that we cannot accept the broad scope of exclusion because to do so commits us to the claim that authorities care about how we are motivated. If I am commanded not to murder, it shouldn't matter whether I don't murder because I was commanded to or because I simply think murdering is wrong. But if the scope of exclusion includes reasons in favor of the command, like the fact that murder is wrong, then I disobey when I am motivated not to murder by the wrongness of murder, and this seems mistaken. All of this assumes the motivational interpretation (as it should, since that is the interpretation Raz endorses and the argument is intended as a critique of Raz). If we switch to the justificatory interpretation, the sharp point of this horn fades away. If the justificatory force of all the reasons in the authority's jurisdiction is excluded, the subject cannot justify her action by appeal to these reasons. But what she can be actually psychologically motivated by is left untouched and there is no dissonance with our

understanding of authority. So the first horn, the broad scope of exclusion, spears the motivational interpretation but not the justificatory interpretation.

On the second horn of the dilemma, we assume the narrow scope of exclusion and the problem is double-counting. If the fact that going to the gym has physical benefits is allowed to count for me in addition to the fact of the command, then when I weigh my reasons it may improperly tip the balance because the authority already took account of its force when coming to her decision. All of this assumes the justificatory interpretation; it assumes that what is affected is the actual force of the reasons and the role they can play in the subject's deliberation, not what the subject may happen to actually be motivated by. The whole metaphor of double-counting is a metaphor of weight and justification, not motivation.

But unlike in the first case, switching interpretations does not simply resolve the conflict. If we assume the motivational interpretation, then the weight or force of any reason is not excluded, only the possibility of being motivated by that reason. This means that the excluded reasons to disobey still have force. The problem of double-counting thus arises again, although in slightly modified form. The problem is not that the subject will double-count and thus systematically favor doing as commanded. It's that, because exclusion doesn't affect weight at all, the effect of counting the authority's judgment will be to double-count all the preexisting reasons, both pro- and con-. This will give those reasons additional and inappropriate weight when compared against reasons that fall outside the authority's jurisdiction and so undermines the good functioning of the subject's deliberation. This problem will arise for the motivational interpretation regardless of the

scope of exclusion because it is a problem about weights and on the motivational interpretation exclusion does not affect weight at all.

So if we follow Raz and accept the motivational interpretation of exclusionary reasons, we run into a dilemma and are either forced to contradict our best understanding of authority or accept the problem of double-counting. I'm not sure if double-counting is very problematic, but we don't need to worry about Raz's response here. If we accept the justificatory interpretation, as I've argued we should, then double-counting in conjunction with the narrow scope of exclusion is still a worry. But the first horn of the dilemma is no longer problematic because on the justificatory interpretation subjects can be motivated however they please, as long as they do as commanded. In response to Essert's argument, then, we should endorse the justificatory interpretation and the broad scope of exclusion.⁴²

The other argument that claims the idea of exclusionary reasons is conceptually incoherent comes from Heidi Hurd. Hurd frames the problem as a problem of rationality. The service conception is attractive in large part because it shows how the space of subjects' reasons is affected so that they can rationally act on the basis of the command. Hurd claims that exclusion would force the subject to act irrationally, so Raz's purported solution to the irrationality of deference to authorities is undermined and his account of authority vitiated. While I do not primarily think of Raz's argument in this way, it is interesting to see why Hurd's argument fails.

Hurd begins by noting that Raz often says authorities require the surrender of personal judgment, and do so by excluding some reasons from the deliberative process.

⁴² This seems correct under the justificatory interpretation because the whole idea of authority is that the force of the authority's command is a function of the force of the underlying reasons, both those reasons in favor of doing as she commands and those against. It seems apt, then, that the force of the command replaces all of them, not just some.

But excluding those reasons leaves the subject in a bind because authority must be evaluated by the subject. Here's the thought: a purported authority commands me to take some act. If it is the command of a genuine authority, I am *ipso facto* bound to obey. In order to know if the authority is genuine, however, I must know whether it fulfills the normal justification thesis, whether the command falls within the authority's jurisdiction, and so on. Investigating these issues requires me to deliberate about the first-order reasons.

Take fulfilling the normal justification thesis: before I can rationally submit to a purported authority, I need to know whether the authority's judgment will help me better conform with the preexisting reasons. In order to know this, however, I need to examine the balance of reasons after I have already been commanded (so know what the authority's judgment is). Otherwise I may follow a command that doesn't help me better conform. But examining the balance of reasons is precisely what I have been excluded from doing. I cannot rationally submit until I know whether the authority's command fulfills the normal justification thesis, but I cannot know whether the authority's command fulfills the normal justification thesis without violating exclusion, so I cannot rationally obey an authority. Hurd pushes this same argument against knowing whether the authority's command falls within her jurisdiction and whether the command runs into any unstated exceptions, for example egregious injustices. On several fronts, then, exclusion prevents subjects from rationally obeying the commands of authorities, and the rational dilemma of authority still holds.

There are two major problems with this argument against exclusionary reasons. The first problem is that, as I have noted, Raz explicitly denies that exclusion has anything to do

with deliberation or what the subject is allowed to think about. Hurd attempts to escape this problem in a footnote:

While, as Raz's point suggests, a commander's order may not, from the commander's point of view, preempt a subject's own deliberation concerning the merits of the order, that order may well preempt such deliberation from the recipient's point of view, since if that recipient is barred from acting on the results of her deliberations, she may wisely find it either pointless or risky to engage in such deliberation.⁴³

This concern seems misplaced to me. The reasons not to deliberate that the recipient responds to are purely pragmatic. When I receive a piece of advice, there is no exclusionary reason present. Still, there may be plenty of pragmatic reasons for me not to engage in any further deliberation, for example if it is likely to be pointless. The fact that I can have these same reasons to cut off deliberation in either the advice or the authority case shows that "preemption" is here a feature of the recipient's context, not any effect the authority had on her. The preemption that doesn't happen, the commander's preemption, is the principled kind of preemption that we should worry about. If the subject finds some pragmatic reasons to deliberate, maybe she just finds idle speculation entertaining, then she is perfectly free to deliberate. There's nothing preventing the subject from deliberating, so she can look at the reasons and assess, for example, whether the authority fulfills the normal justification thesis if she likes.

The second problem with Hurd's argument is that it misinterprets the normal justification thesis and the demands of rationality. Consider the following passage:

If it is rational to abide by the laws enacted by a practical authority only if that authority is legitimate (premise (4") above), and if an authority is legitimate only if its laws *better cohere with* the balance of content-dependent reasons for action than do the judgments of those for whom it is an authority (the normal justification thesis), then it must be the case that in

⁴³ Hurd, "Challenging Authority, p. 1626, n. 27.

order to judge whether indeed an authority is acting legitimately one must oneself balance the reasons for action *in each case* in which a law applies so as to police the ability of the claimed authority to order action in conformity with that balance. The ability of a claimed authority to achieve action which in the long run better accords with the balance of reasons can only be measured if, *at each decision*, one judges for oneself the reasons for action, and compares one's judgment with that reached by the authority. Without engaging in such long-term comparisons, one has no basis for thinking the claimed authority legitimate, and hence no rational foundation for abiding by that authority's will. One lacks, that is, any foundation for thinking that Raz's normal justification thesis applies. Yet since the attribution of practical authority to governmental institutions bars one from engaging in just the sort of comparisons that practical rationality depends upon, that attribution cannot be rationally defended.⁴⁴

I have italicized the problematic parts of this argument. Hurd reads the normal justification thesis as addressing whether obeying a particular command will help the subject better conform with reason *on that occasion*. Here is Raz's formulation of the normal justification thesis, which is importantly different:

[T]he normal way to establish that a person has authority over another person involves showing that the alleged subject is *likely* better to comply with reasons which apply to him (other than the alleged authoritative directives) if he accepts the directives of the alleged authority as authoritatively binding and tries to follow them, rather than by trying to follow the reasons which apply to him directly.⁴⁵

The key word is 'likely'. To know that an agent fulfills the normal justification thesis and so is an authority, one need only know whether the agent's judgment is more *likely* to better cohere with the preexisting reasons.⁴⁶ And this can be known without looking at each case, and indeed can be known prior to knowing what the authority's particular judgment is in this case. Once I have a sufficiently robust track record for the authority, I can infer that her judgment will continue to be better than mine and rationally defer my judgment to her.

⁴⁴ Hurd, "Challenging Authority", p. 1633. My emphasis.

⁴⁵ Raz, MF, p. 53. My emphasis.

⁴⁶ Cf. Raz, MF, p. 47: authority is justified "not by assuming that they always succeed in acting in the ideal way, but on the ground that they do so often enough to justify their power".

This is true of all sorts of cases. I'm standing outside and it feels 65° to me, but then I look at the thermometer and it reads 62°. I defer my judgment about temperature to the thermometer and believe it's 62°. Why is it rational for me to do this? It's not because I then checked another thermometer and confirmed that it was 62° or anything of the sort. It's because I know the thermometer has been reliable in the past and I have good reasons for thinking that reliability will continue into the future. In the thermometer case these reasons have to do with its physical construction and durability. These reasons can give me sufficient confidence in the better reliability of the thermometer for me to rationally defer my judgments about temperature to it. To claim otherwise is to endorse a very implausible view of the rational agent as someone who always needs direct evidence. If this were the case, the vast majority of our beliefs and acts would be irrational.

Hurd explicitly denies that I can make inferences about the future reliability of sources, but her reason is based on the misreading of the normal justification thesis.⁴⁷ Once we realize that the normal justification thesis, and so genuine authority, is about tendencies and not actual conformity in every case, the rational basis for deference is much easier to attain. Indeed, an account of authority must leave open the possibility of mistaken (i.e., non-conforming) yet binding commands⁴⁸ and it is only because the normal justification thesis is about tendencies and not instances that this possibility is open.

If Hurd were correct that rational deference requires examining each case of deference one by one, more than just exclusionary reasons and Razian practical authority would fall. So would epistemic authority, for deference to epistemic authorities is based on the fact that deferring to them leads to better conformity with the preexisting reasons to

⁴⁷ Hurd, "Challenging Authority", p. 1635.

⁴⁸ As Hurd implicitly admits, *ibid.*, p. 1636.

believe. In the practical authority case, the problem for deliberation was that exclusion purported to cut deliberation off. In the theoretical authority case, the problem for deliberation is most often that the layperson *can't do it*. The expert has much more knowledge and experience, and assessing the claims of an expert is nearly always impossible for the layperson. I cannot, for example, reliably judge whether the judgments of a physicist are true because I simply don't have the knowledge or skill to assess all the relevant facts and theories in physics. If a barrier to deliberation is a barrier to deference, then deference to an expert like the physicist is always irrational, but this is clearly not the case.

Hurd's argument is correct that it will often be quite difficult for subjects to determine whether they are obligated to obey a particular command. They need to know that the authority is legitimate, that the command falls within the scope of her jurisdiction, that there are no decisive reasons to disobey from outside her jurisdiction, and so on. But this is just a feature of living and acting in the actual, messy, complicated world. We constantly have conflicting demands placed on us, reasons to act in a wide variety of incompatible ways. This makes acting well difficult, but it is a general feature of life in complex contexts that explains the difficulty, not any particular problem with commanding and certainly not any problem with exclusion. The practical problems Hurd highlights are definitely an issue, but as Alvin Goldman has argued, this difficulty does not preclude the many indirect ways laypeople have to identify experts.⁴⁹ While a difficult practical task, it is certainly not insurmountable and in fact is something we do on a regular basis.

⁴⁹ Goldman, "Experts: Which Ones Should You Trust?"