

Cardozo Law School

Jacob Burns Institute for Advanced Legal Studies

Public Law and Legal Theory

Working Paper 024

Stanford/Yale Junior Faculty Forum

Research Paper 00-05

2000

Authority

Scott J. Shapiro

This paper can be downloaded without charge from the
Social Science Research Network Electronic Paper Collection:

http://papers.ssrn.com/paper.taf?abstract_id=233830

AUTHORITY

*By Scott J. Shapiro*¹

The Babylonian Talmud tells of a dispute between Rabbi Eliezer and the other rabbis over the ritual cleanliness of some pieces of tile.² Rabbi Eliezer argued at length that the tiles were clean, but failed to persuade his colleagues. Having tried every conceivable argument, Rabbi Eliezer said to the rabbis: "If the Law agrees with me, let this carob-tree prove it!" whereupon the carob-tree was torn from its roots and landed 100 cubits away.³ The rabbis answered: "No proof can be brought from a carob-tree." Rabbi Eliezer tried again: "If the Law agrees with me, let the stream of water prove it!" and behold, the stream obliged and reversed direction. Undaunted, the rabbis responded: "No proof can be brought from a stream of water." Rabbi Eliezer came back a third time: "If the Law agrees with me, let the walls of the schoolhouse prove it." As before, Rabbi Eliezer's request was granted and the wall started to fall. But Rabbi Joshua rebuked the wall, saying: "When the scholars are engaged in a legal dispute, what have ye to interfere?"

At his wits end, Rabbi Eliezer exclaimed: "If the Law agrees with me, let it be proved from Heaven!" whereupon a voice from heaven cried out: "Why do ye dispute with Rabbi Eliezer, seeing that in all matters the Law agrees with him!" But Rabbi Joshua replied: "It is not in Heaven".⁴ Rabbi Jeremiah explained this to mean that once God gave the Torah to Moses on Mount Sinai, Heavenly voices no longer have any say in legal matters. The Torah confers on the rabbis the exclusive authority to decide

¹ Associate Professor of Law, Benjamin N. Cardozo School of Law, Yeshiva University. Dis claimer: In keeping with the spirit of the Forum, this paper is still very much a work in progress. Please do not cite or quote without the author's permission.

² Baba Metzia, 89 a,b.

³ For some reason, the Talmud adds that others claim that the Carob tree landed 1000 cubits away.

⁴ Deuteronomy 30: 12

legal questions and requires that internal disputes be resolved by majority vote. After Rabbi Eliezer refused to back down, the story ends with his excommunication.

This story brings out, in a particularly striking way, the paradoxical nature of authority. Authorities claim a right of immense power, one, it would seem, that they could not possibly possess. Authorities claim the right to impose their beliefs on others regardless of whether their beliefs are correct. In doing so, they appear to place themselves above the truth – their right does not seem to depend on their being right. In the dispute between Eliezer and the rabbis, the rabbis had incontrovertible proof that they were wrong and nevertheless continued to demand that Eliezer accept their interpretation of the law. This sounded, at least to Eliezer, as though the rabbis were arrogantly asserting a power greater than that of Heaven. If the rabbis are wrong, he reasoned, they should submit to God, not require submission from others.

Eliezer may be recorded history's first philosophical anarchist. Philosophical anarchists argue that no claim to authority can be vindicated. How can someone have the right, they wonder, to force another to do something wrong? The obligation to act correctly should always trump the obligation to act otherwise. To the philosophical anarchist, "legitimate authority" is a contradiction in terms.

Philosophical anarchists delight in pointing out that the claims of authority are problematic even in situations when those in power are right. For when someone in authority commands another to act as they should act, their directives are redundant. They would not harm, but they would not help either. Thus, even if the rabbis had relented and declared the tiles ritually clean, Eliezer would have had no reason to submit to their authority. Eliezer should accept that the tiles are ritually clean because they are ritually clean, not because the rabbis said they are.

The challenge posed by the philosophical anarchists turns out to be as powerful as it is simple:

when authorities are wrong, they cannot have the power to obligate others—when they are right, their power to obligate is meaningless. It would seem that the institution of authority is either pernicious or otiose.

This argument is so powerful, in fact, that it should make one suspicious. For if the argument is valid, then those who believe in authority are not just wrong—they are incoherent. This conclusion seems too strong, however: those who believe that they are obligated to obey are not believing nonsense. In the end, no such obligation may actually exist; but it seems conceivable that such an obligation could exist.

Most theorists writing today assume that common sense is correct and that the anarchist challenge can be met. They disagree, however, on how to meet it. In this essay, we will discuss the many “solutions” that have been offered on authority’s behalf. The responses fall roughly into one of two groups: those who believe that problems arise due to certain naïve views about the nature of authority and rationality and that revision in our understanding is required, and those who maintain that the puzzle can be unraveled without any radical changes.

After a discussion of the paradox (or as we shall see paradoxes), we will examine two revisionist strategies. The first approach denies that legitimate authorities have the right to impose obligations when they are wrong. Indeed, it denies that legitimate authorities have the right to impose obligations at all. The directives of legitimate authorities are treated either as justified threats backed by sanctions or pieces of expert advice. The paradoxes, on this view, simply dissolve – because authorities never possess the power to impose obligations, *a fortiori* they never possess the power to impose obligations when they are wrong.

The second approach we will explore accepts the basic premise of the paradoxes, namely, that

legitimate authorities have the power to obligate even when they are wrong. But it argues that the paradoxes arise because that premise is conjoined with a tacit, but false, assumption about the nature of rationality. According to this assumption, agents must always act on the merits of the case at hand. Instead, this approach maintains that agents can, under certain conditions, have reasons to ignore the desirable or undesirable properties of actions. There is nothing paradoxical, therefore, about requiring obedience to mistaken directives, because the directives are reasons not to act on the reasons that make the directives mistaken.

Unsurprisingly, we will see that these responses bring additional problems in their wake. For if we deny that authorities ever possess the power to obligate, can we make sense of social institutions such as the Law that contain copious amounts of prohibitions, requirements, permissions, rights and powers? Likewise, is it even coherent to claim that we can have reason to ignore reasons for action?

In an effort to avoid these problems, others have attempted to solve the paradoxes without revisionism. While they concede that legitimate authorities have the power to obligate even when they are wrong, they argue that standard theories of rationality and morality can accommodate this right. Their strategy centers on the claim that authoritative directives of legitimate authority are efficient decision-making tools. By guiding conduct by legitimate authority, subjects are more likely to choose the right results than the wrong results. But, these theorists argue, subjects must take the bad with the good – subjects can benefit from the right results only if they choose the wrong results as well. The rationality of obeying a mistaken directive is, therefore, no more paradoxical than the rationality of paying a price for a gamble, when the price is less than or equal to the gamble's expected value.

While this approach is appealing, I will argue that it ultimately will not do. The paradoxes of authority cannot be solved within standard theories of rationality and morality – some revisions are necessary. Which revisions are necessary, I will claim, depends on one's underlying theory of legitimacy.

For accounts that tie the legitimacy of authority to its ability to provide instrumentally valuable directives, I will suggest that the standard account of authority's effect on practical reasoning be modified. Instead of seeing authoritative directives as instruments that willing subjects use to make decisions, they ought to be understood as causal constraints on action. Those who obey directives in order to instrumentally benefit from them do not choose to obey – having submitted to authority, disobedience is no longer an option. Authoritative directives can be justified in instrumental terms when, and only when, they forestall decision-making.

For accounts of authority that tie legitimacy to the moral obligation to respect collective decision-making procedures, I will suggest that we modify our views about the nature of moral autonomy. In certain circumstances, the fact that another has demanded that we act can indeed give us a reason to act. Rather than a violation of autonomy, obedience can actually show due respect for the value of autonomy.

In this essay, I will attempt to justify these assertions and to demonstrate how their acceptance solves the paradoxes of authority within the different frameworks of legitimate authority. While these revisions may be somewhat drastic, I will argue that they constitute the best response the philosophical anarchist's challenge.

I. The Paradoxes of Authority

A. Authority and Autonomy

In his In Defense of Anarchism, Robert Paul Wolff argued that legitimate authority and moral autonomy are logically incompatible.⁵ His discussion is worth examining in detail, not only because it is the *locus classicus* for the philosophical anarchist attack on authority, but also because it contains a more subtle analysis of the concept of authority than many of authority's defenders provide.

Wolff begins his discussion by distinguishing between power and authority.⁶ To have power is to have the ability to compel others to do as one wants. To have authority is to have the right to rule. A gunman has power, but he does not have authority. He can coerce his victim to cooperate by threat of force, but he is unable to impose the obligation to comply. Unlike authority, power cannot be honored in the breach: one can cheat at one's taxes, but one cannot cheat a thief.

As Wolff points out, someone can have authority in one of two senses.⁷ One can have authority by possessing the moral right to rule. The exercise of such a right, if it exists, genuinely gives rise to moral obligations to obey. A ruler can, therefore, claim authority and fail to have it in this sense. The Supreme Soviet Legislature claimed the authority to rule the Soviet Union, but it lacked the moral right to do so. It lacked legitimate, or *de jure*, authority. The Supreme Soviet, did enjoy a measure of acceptance by many of the Soviet people. They believed that the Supreme Soviet possessed *de jure* authority, although

⁵ See Robert Paul Wolff, In Defense of Anarchy (1970). The relevant sections of Wolff's monograph are excerpted as "The Conflict between Authority and Autonomy" in Authority (Joseph Raz ed., 1990). All citations in this essay will be made with reference to this excerpt.

⁶ Wolff, *supra* note 5, 20.

⁷ *Id.* at 21.

they were mistaken—it merely had *de facto* authority.

Wolff is primarily interested in the phenomenon of *de jure* authority. This is so for two reasons. First, the concept of *de facto* authority logically presupposes the concept of *de jure* authority. One who possesses *de facto* authority is someone whose claims to *de jure* authority are believed by a significant portion of its subjects. Second, it is uncontested that *de facto* authority exists. The philosophical anarchist is interested in whether *de jure* authority exists, because she wants to show that the moral obligation to obey the law can never obtain.

To have the right to rule, according to Wolff, is to have the right to be obeyed. To obey an authoritative command is to perform the act commanded for the reason that it was commanded. Commands, therefore, differ from arguments.⁸ An argument is meant to persuade. It attempts to convince the person that they ought to act in a certain way and it does this by presenting to the interlocutor the reasons that make the action worthy. One who issues a command, on the other hand, does not intend to convince the subject of the wisdom of his order. The commander does not give reasons why the action commanded is worthy of obedience, but rather demands that his command be taken as a conclusive reason for obedience.

It is possible, therefore, for someone to conform to a command without obeying it.⁹ This happens when the command makes the subject aware that he has reasons for performing the act commanded and acts for these reasons, rather than because of the command. To acknowledge someone's claim to authority, according to Wolff, is to recognize that their right resides in their person.¹⁰ They possess their

⁸ *Id.* at 22.

⁹ *Id.*

¹⁰ *Id.* But see discussion in Section III B., *infra*.

power in virtue of who they are, rather than in virtue of what they command.

Having set out his conception of authority, Wolff proceeds to present his account of moral autonomy. For Wolff, an autonomous person is not someone who is merely responsible for her actions. Rather, such a person also takes responsibility for her actions.¹¹ A person takes responsibility whenever she attempts to determine what she ought to do. An autonomous agent, according to Wolff, is a deliberating agent.

Unlike many who have seen autonomy as a necessary condition for moral responsibility, or as the capacity to choose, Wolff treats it as an independent moral duty. Every person is charged with examining every aspect of his moral life: he must constantly gather new information, scrutinize his motives, critique his desires and evaluate his options in light of this reflection. One who acts without assessing the merits of so acting fails to take responsibility for his actions, and to this extent, is violating their duty to act autonomously.¹²

It follows from Wolff's definition of moral autonomy that no one can obey authority and remain autonomous. A person obeys a command when he conforms for the reason that another has so commanded. An autonomous person, however, never acts for the reason that another has so commanded. He acts only when he is convinced that, on the merits, action is appropriate. Hence, an autonomous agent can never submit to another's authority. As Wolff puts it: "The defining mark of the state is authority, the right to be ruled. The primary obligation of man is autonomy, the refusal to be ruled."¹³

¹¹ *Id.* at 25.

¹² *Id.* at 28-29.

¹³ *Id.* at 29.

Although the autonomous agent cannot obey authority, Wolff is quick to add that he does not necessarily disobey authority.¹⁴ If the autonomous agent thinks there are good moral reasons to pay taxes, then he will believe that he should pay his taxes. But that person does not accept the obligation because the law requires him to pay his taxes. He believes that he should pay his taxes because he believes this to be the right thing to do independent of the law's demands.

B. Preliminary Assessment

Wolff's argument appears to be valid: given his premises, his conclusion seems to follow. However, his premises are dubious. For example, it is not obvious that one should think of autonomy as a duty. To whom do we owe this duty? It is strange to think that I am morally bound to act autonomously for another's sake. Why would anyone care why I act correctly, as long as I act correctly?

Even if one does accept that there is a duty to act autonomously, it is doubtful that Wolff's formulation should be accepted. Why must a person deliberate about every moral action? Shouldn't he defer to another's judgment when that judgment is better than his? The idea that a person must weigh the balance of reasons every time a moral decision arises is not only dangerous in cases of informational asymmetries or cognitive disabilities but is also terribly wasteful. Surely one could focus his energies in a more productive use than constant deliberation.

To be sure, there is such a thing as over-reliance on authority. To cede too much decision-making to others is both foolhardy and morally irresponsible. Moreover, the more one depends on

another's judgment, the greater the chance that one will lose the ability to make judgments for oneself and the more vulnerable one becomes to manipulation. Finally, the faculties of judgment and self-reflection are distinctively human capabilities, the exercises of which contribute in an essentially way to human flourishing. To sacrifice them is, in some real sense, to forfeit one's humanity. This "dehumanizing" effect of authority especially concerned William Godwin, the first "modern" philosophical anarchist.

Man is the ornament of the universe, only in proportion as he consults his judgment. . . . But, where I make the voluntary surrender of my understanding, and commit my conscience to another man's keeping, the consequence is clear. I then become the most mischievous and pernicious of animals. I annihilate my individuality as a man, and dispose my force as an animal to him among my neighbors, who shall happen to excel in imposture and artifice, and to be least under restraint from the scruples of integrity and justice.¹⁵

While the dangers of reliance on authority are real, it is important not to exaggerate them. The world is simply too complex for anyone to live one's life completely unaided by experts of one kind or another. Even Wolff admits that "[t]here are great, perhaps insurmountable, obstacles to the achievement of a complete and rational autonomy in the modern world."¹⁶ Complete autonomy, in Wolff's sense, is simply not an option. If authority is inconsistent with autonomy, then so much the worse for autonomy.

It is interesting to note that Kant himself did not see a clash between authority and autonomy. He famously argued that those subject to authority ought to question its demands, but this public use of reason should not prevent them from acting on them. Enlightenment is precluded both when authority demands blind obedience and when subjects do not respond with unconditional compliance. "The citizen cannot

¹⁴ *Id.*

¹⁵ William Godwin, ENQUIRY CONCERNING POLITICAL JUSTICE 122 (K. Carter ed., 1971).

¹⁶ *Id.*

refuse to pay the taxes imposed upon him; presumptuous criticism of such taxes, where someone is called upon to pay them, may be punished as an outrage which could lead to general insubordination. Nonetheless, the same citizen does not contravene his civil obligations if, as a learned individual, he publicly voices his thoughts on the impropriety or even injustice of such fiscal measures."¹⁷

Wolff's formulation of the anarchist's challenge is unconvincing because his understanding of autonomy is implausible. We should, however, be careful not to dismiss Wolff's argument too quickly, for on any credible conception of autonomy, the tension between it and authority is hard to ignore. After all, "autonomy" literally means "self law-giving." The autonomous person does not act simply because another has told him to do so—he acts only when convinced that action is appropriate. To be autonomous, in other words, involves taking oneself as the ultimate authority on moral questions. This commitment seems to leave no logical space for external authorities to occupy. As the proverb goes, one cannot serve two masters.

With this in mind, I think it is possible to give a more charitable reading to Wolff's objections. We should first distinguish, in a way that Wolff fails to do, between two different features of authoritative directives. We can say, following H.L.A. Hart, that authoritative directives are both "peremptory" and "content-independent" reasons for action.¹⁸ A "peremptory" reason is a reason that preempts deliberation. A command is a peremptory reason, in that once the command has been issued, the time for deliberation comes to an end and the time for action begins. The obeying agent no longer considers the merits of following the command and simply acts as he is directed.

¹⁷ Kant, "An answer to the question: What is enlightenment?" in Kant: Political Writings (H. Reiss ed., 1989).

¹⁸ H.L.A. Hart, "Commands and Authoritative Reasons" in Essays on Bentham, 253.(1982).

A directive is a “content-independent” reason when it gives an agent a reason to comply irrespective of whether the agent has a reason to act on its content. The fact that such a directive requires that an act be done is, by itself, a reason to perform that act. One who obeys a command, therefore, treats the command as a content-independent reason, because he complies for the reason that he was commanded, not because he has reasons to act on the content of the command. For example, if Jim takes out the garbage because his father commanded him to do so, then he is treating the command as a content-independent reason.

Content-independent reasons for complying with a directive should be contrasted with “content-dependent” reasons. A content-dependent reason is a reason for conforming to a directive because the directive has a certain content. If the garbage smells, Jim will have a reason for taking out the garbage that is independent of the fact that his father commanded him to do so. By taking out the garbage, he will have removed an unpleasant odor from the house. Jim, therefore, has two reasons to listen to his father’s command: the command is a content-independent reason, while the unpleasant odor is a content-dependent reason.

Although Wolff appears to object solely to the peremptory nature of authority, I think that it is the combination of peremptoriness and content-independence that offends him. Authority and autonomy clash not simply because one who obeys does not deliberate. The problem is also that such a person believes that the fact that he was ordered to act in a certain way gives him a reason to so act. He takes the will of another as his reason, indeed the only reason, rather than the merits of the case at hand.¹⁹ Such

¹⁹ See, e.g., Wolff, *supra* note 5, 26 (“The autonomous man, insofar as he is autonomous, is not subject to the will of another. He may do what another tells him, but not *because* he had been told to do it.” (emphasis in original)).

a person, therefore, will think that he has a ready defense to any charge of improper behavior. While the person will agree that he performed a wrongful act, he will plead that the reasonableness of his actions must be viewed in a content-independent manner: whether he had reason to follow orders cannot be judged based on the content of those orders. It is the fact that he was commanded to act, rather than what he was commanded to do, which gave him a conclusive reason to do as he did.

An autonomous person, by contrast, never treats a command as a content-independent and peremptory [hereinafter “CIP”] reason for action. The demands of authority mean nothing to the autonomous agent, for such a person never allows his will to be determined by the will of another. She cares solely about the act commanded, not the command itself, and will acquiesce only when convinced that there are good reasons to act on the content of the command. According to this interpretation, autonomy and authority are incompatible because obedience to authority requires acting on CIP reasons, whereas the autonomous person does not acknowledge the existence of such reasons.

One benefit of seeing Wolff’s argument in this way is that our previous objections are no longer sufficient to meet his challenge. Autonomy is not conceived as a separate duty that morality imposes upon us and that we owe to others. To say that everyone should act in a morally autonomous manner is to make a claim about the space of reasons. Autonomous agents are those who recognize that the only reasons that exist are either content-dependent or non-peremptory ones. Moral autonomy is important because it is important that people act on reasons and not act on non-reasons.

Moreover, on this account, reliance on experts does not necessarily lead to heteronomy. While expert advice is a CIP reason for believing that the expert is correct – one believes what the expert says simply because the expert has said it – the purpose of giving advice is to alert the advisee that the

recommended course is supported by the balance of content-dependent reasons. When the advisee draws this inference and acts on it, the agent will be acting for content-dependent reasons, even if he does not know what they are.

Most importantly, this interpretation shows that the philosophical anarchist's anxiety about authority is not frivolous: their worry is that people will treat authoritative directives as CIP reasons for action and, in so doing, fail to take the appropriate responsibility for their actions. They will attempt to justify their conduct by pleading that they were "just following orders." This type of defense not only seems cowardly, but strictly speaking irrelevant. How can an act be made acceptable simply because someone else says that it is acceptable? Authorities may have the power to change positive law, but no one (maybe not even God) has the ability to change the moral law. As Godwin put the point: "The most crowded forum, or the most venerable senate, cannot make one proposition a rule of justice that was not substantially so, previously to their decision."²⁰

Yet, the philosophical anarchist reminds us, without admitting that Nuremberg defenses are sometimes good justifications, it is hard to see how *de jure* authority is possible. The legitimacy of authority stands or falls on whether a subject can justify his actions by pleading that he was "just following orders." Legitimate authority is possible, in other words, only when CIP reasons are possible.

C. Authority and Rationality

It is sometimes thought that Wolff's challenge to authority is merely a special case of a more

²⁰ Godwin, *supra* note 15, 88.

general paradox, one that purports to show the incompatibility of authority and rationality. The general argument is familiar: Consider any directive issued by an authority and any action required by that directive. Either the balance of reasons supports that action or it does not. If the balance of reasons supports the action, an agent should conform to the directive, but not because conformity is required by the directive, rather because agents should always act according to the balance of reasons. On the other hand, if the balance of reasons does not support the action, then an agent should not conform to the directive because agents should never act against the balance of reasons. It would seem, therefore, that authoritative directives can never be reasons for action – if a directive gave the right result, the directive would be irrelevant; if the directive gave the wrong result, then the obedience to the directive would be unreasonable.

Since authoritative directives can never be reasons for action, it follows that rational agents can never obey authority. The proof: Rational agents always aim to act on undefeated reasons and act in accordance with that aim. If an agent were to obey an authority, they would either have to believe that they had an undefeated reason to obey or believed that they didn't have an undefeated reason but would have obeyed anyway. If the former were true, then the agent would have irrational beliefs, given that according to the first argument, authoritative directives can never be reasons for action. If the latter were true, then the agent would not be acting in accordance with the aim of acting on undefeated reasons. Hence, it seems that rational agents can never obey authority.

If the above arguments are sound, it would follow that moral agents can never rationally guide their conduct by authoritative directives. Since morality requires that agents act on the balance of moral reasons, obedience to authority can never be rationally justified for moral agents: whenever a directive

required an action supported by the balance of moral reasons, that directive would be morally irrelevant; otherwise, it would be morally pernicious. Authoritative directives can never be moral reasons for action and, hence, it would be irrational for any moral agent to obey authority.

Such “derivative” arguments are possible because rationality is essentially a formal ideal. Rationality does not mandate conformity to any particular standard – it simply requires that an agent live up to the standards that the agent judges he should live up to. The paradox attempts to show the incompatibility of rationality and authority by demonstrating that authoritative directives will clash with any normative standard: either the directive in question conforms to the given standard, in which case it is redundant, or it conflicts with the standard, in which case the standard requires nonconformity. To generate a contradiction between authority and any specific normative standard, one need only plug the standard into the equation and out will pop the desired *reductio*. While Wolff’s challenge appears to be such a derivative argument, it is important to see that it is not. This becomes evident when it is noted that the concept of rationality is at right angles to that of autonomy. To be rational is to aim to act on undefeated reasons and to act in accordance with that aim. To be autonomous, by contrast, is to aim to act on non-CIP reasons and to act in accordance with that aim. It does not follow, therefore, that a rational agent is an autonomous agent. If an agent believes that he has an undefeated CIP reason for action, then he will be acting rationally but not autonomously if he acts for this reason. Conversely, autonomous agents are not necessarily rational. If an agent acts on a content-dependent reason that, by his own lights, is defeated, then he will be acting autonomously but irrationally.

It should come as no surprise, therefore, that one can challenge the rationality of authority and not its effects on autonomy. The paradox of authority and rationality attempts to show the impossibility of

having an undefeated reason to obey authority and, hence, the irrationality of believing that one can have such a reason. It does not attack the content-independent and/or peremptory nature of authoritative directives. Likewise, one can object to authority because it engenders heteronomy, not irrationality. The problem with obedience, according to our interpretation of Wolff's challenge, is that authoritative directives are not CIP reasons, not that it would be incoherent to believe that they are undefeated reasons.

Given that these critiques differ from each other, one should not expect that the solution to one challenge will constitute an effective reply to the other. To see this, consider the following response to the paradox of authority and rationality: The social contract theory, the response begins, is a coherent theory of political obligation. A rational agent might regard it as true, even if it is not true. Assume, then, that an agent accepts the social contract theory as true. According to this agent, someone possesses legitimate authority over another when the latter has consented to be ruled by the former. Because the consent generates a promissory obligation to abide by the demands of the authority, any directive issued gives subjects who have consented a reason to act in accordance with it. Assume that this person consents to be governed by an authority. He will now regard any directive issued as a reason for action. Consequently, it might be rational for him to comply with a command whose content, by his own lights, is not supported by the balance of content-dependent reasons. From that agent's perspective, even though the balance of content-dependent reasons would not support conformity, the balance of all reasons – content-dependent reasons as well as the content-independent reason – might tilt in the direction of obedience. In this way, it may be rational to obey authority even when they are wrong about the content of their directives.

This response, however, will not work against Wolff's challenge. Wolff's argument, as we have

seen, is predicated on the idea that there are no such things as CIP reasons for action. The above response would, therefore, beg the question. After all, consent itself purports to be a CIP reason for action. Subjects who consent to be governed by an authority are obligated, under the social contract theory, simply in virtue of their consent. One cannot show how a CIP reason is possible by producing another (alleged) CIP reason. One must first establish that my will can give me a reason to act against the balance of reasons. But if authorities lack the power to change the moral law, how can I have the power to do so?

I don't mean to imply that Wolff's autonomy paradox is harder to answer than the rationality paradox. They are simply different critiques and, as such, each may require different solutions. Unfortunately, those who respond to philosophical anarchism do not always make it clear to which paradox they are responding. We will try to rectify this by treating the two basic paradoxes separately and asking, for every response by the defenders of authority, whether the solution is adequate to either or both.

II. Weakening Authority

Whenever faced with the clash of two concepts, one can always try to relieve the tension by weakening the formulation of one of the concepts involved. In the case of the anarchists' paradoxes, the most obvious candidate is the one given to authority. First, because the concept of authority appears in both of the basic paradoxes, one might be able to kill two birds with one stone. Second, and more importantly, the formulations of rationality and autonomy seem innocuous enough, at least as compared to

that of authority. As we have mentioned, to be rational is to aim to act on the balance of reasons. Rationality simply imposes the obligation on every agent to do whatever he believes he is supposed to do.

Likewise, autonomy has been characterized as requiring that agents stick to the merits of the case at hand and not simply act because another has told one to do so. The plausibility of these ideas is readily apparent.

By contrast, the standard characterization of authority wears paradox on its sleeve. To possess legitimate authority, one might recall, is to have the right to rule. The right to rule implies the right to be obeyed. To have the right to be obeyed is to have the power to impose obligations irrespective of content. Those who possess legitimate authority, therefore, have the power to obligate others to obey even when their directives contain the wrong content. To say the very least, the power to obligate independent of content seems odd. As Godwin remarked: “There cannot be a more absurd proposition, than that which affirms the right of doing wrong.”²¹

Weakening the standard characterization of authority, therefore, seems like a promising strategy. If it can be shown that legitimate authorities never claim, and do not have, the power to obligate, then the paradoxes could be completely sidestepped. Rationality would not clash with authority because authorities would never require agents to act against the balance of reasons. Indeed, authorities would never require agents to act at all. Likewise, autonomy would be consistent with authority because their pronouncements would not be, nor would anyone claim them to be, content-independent reasons for conformity.

In the next sections we will explore two strategies for severing the connection between legitimate

authority and the right to require conduct in a content-independent manner. The first approach seeks to decouple the right to be obeyed from the right to rule. The right to rule consists in the exclusive privilege that legitimate authorities possess to coerce others to conform to their demands. On this view, the authorization to use force does not imply the power to impose obligations. The second strategy denies that legitimate authorities even have the right to rule, let alone the right to be obeyed. Legitimate authority issue neither directives nor threats. Their pronouncements constitute expert advice: rather than creating obligations, they simply inform us about their existence.

I will argue that neither approach is successful. To weaken the concept of authority in these ways is to emasculate it. Without attributing to legitimate authorities the power to impose obligations, it would be impossible not only to account for many of the claims authorities do make, but also the claims that they need to make in order to render the concept coherent.

A. The Right to Rule

As Hohfeld taught us, the word “right” is ambiguous.²² To say that someone has a right to perform some act with respect to another might, in the first instance, refer to a privilege. A right-holder has a privilege against another to perform an act when the right-holder is not under a duty to that person not to perform that act. For example, I have the privilege against you to enter my home in that I am not under a duty to you not to enter my home.

Alternatively, one might mean by the right-ascription that the right-holder has the power over the

²¹ Godwin, *supra* note 15, 88.

²² Wesley Newcombe Hohfeld, Fundamental Legal Conceptions as Applied in Judicial Reasoning 35 (W.W. Cook ed.,

other person with respect to a certain set of acts. A right-holder has a power over another to perform a certain act when the right-holder has the ability to change the normative relations between the person and the right-holder with respect to that act. To say that I have the power to lease my house to you is to say that I can change your duties to, and privileges against, me vis-à-vis the house. Whereas previously you did not have the privilege against me to enter my house, after the execution of the lease, you are no longer under a duty to me not to enter. My right to lease the house to you, where the right is understood as a power, leads to your right to enter the house, where the right is understood as a privilege.

When we said that to possess legitimate authority is to have the right to rule, should we take this right to be a privilege or a power? Until now, we have been assuming that legitimate authorities possess normative power over their subjects. Robert Ladenson has suggested, however, that the right to rule ought to be understood as the exclusive privilege to rule.²³ Legitimate authorities are those who have the ability to force others to comply with their demands and who are morally permitted (that is, not under a moral duty not) to exercise this ability. On this view, authorities do not issue norms and their pronouncements create no duties to obey. Rather, authoritative directives are threats backed by sanctions. Legitimate authorities differ from gangsters in that the former are morally permitted to issue such threats and punish non-conformity, as opposed to latter who are under a moral duty not to act similarly.

According to Ladenson, someone has legitimate authority whenever they have (non-normative) power to coerce and that power justifies the exercise of that power. Why would the mere possession of

1919).

²³ Robert Ladenson, "A Defense of a Hobbesian Conception of Law," 9 *Philosophy and Public Affairs* (1980), reprinted in *Authority*, supra note 5.

power justify the exercise of the power? Why think that might makes right? Ladenson offers a “Hobbesian” answer.²⁴ If some institution has the power to coerce, then they, and only they, have the ability to solve certain problems, like maintaining social order. Only the strongest in society can keep others from engaging in the internecine battles that threaten to destabilize and destroy society.

Ladenson does not go so far as to claim that those with power are justified in using this power in any manner they choose. The Nazis were not morally permitted to engage in genocide, even though they were justified in their effort to enforce the existing traffic laws.²⁵ Authorities may claim legitimate authority in certain areas but fail to possess it. Someone has supreme legal authority just in case they have the ability to enforce their threats and those threatened are aware of this ability. This does not, all by itself, confer upon them absolute legitimate authority.

The value of Ladenson’s account is that, if true, it provides a simple answer to the paradoxes of authority. It is a consequence of this view, for example, that the guidance of authority may be both rational and relevant. This will occur whenever the balance of reasons is shifted in favor of conformity due to a threatened penalty. Authorities supply a reason to obey that the agent did not have before the authoritative intervention. Through the coercive machinery of the state, authorities can be seen to make differences to the practical reasoning of agents.

Likewise, obedience to authority would not compromise the autonomy of the agent. On Ladenson’s account, authoritative directives are not content-independent reasons.²⁶ Agents never obey

²⁴ *Id.* at 38.

²⁵ *Id.* at 39.

²⁶ Compare this to Hobbes’ own characterization of authoritative directives: “A command is when a man saith do this or do not do this yet without expecting any other reason that the will of him that said it.” That is, Hobbes took commands to be content-independent reasons for action. The only Hobbesian element of Ladenson’s account,

simply because they are told to do so, but only to avoid being punished. Indeed, even Godwin believed that agents are permitted to obey authorities when threatened: “Nothing can be more certain, that an action, suppose of inferior moment or utility, which for its own sake might be right to be performed, it may become my duty to neglect, if I know that by performing it I shall incur the penalty of death.”²⁷

If accepting Ladenson’s account were the only way to overcome the paradoxes, then the defenders of authority should concede defeat. For while the phrase “the right to rule” is equivocal, the law contains many disambiguating expressions that confirm that the right claimed is a normative power. For example, the law imposes numerous duties upon agents to do or forbear from acting in certain ways. The rule against murder does not simply threaten punishment if one commits homicide with malice aforethought. It imposes a duty on people not to murder and prescribes a punishment for violation of this duty. Indeed, as Hart pointed out in his critique of Austin, this rule would be intelligible even without a prescribed sanction. The rule sets a standard that people are expected to follow and against which their conduct is to be evaluated. To attempt to reduce all legal rules to threats backed by sanctions would be to ignore the internal point of view, that is, the point of view of those who see the authority of the law as legitimate, as having the normative power to set standards of acceptable conduct that all are obligated to follow.

Even from the perspective of power politics, Ladenson’s proposal is implausible. As Weber pointed out, political power is in need of legitimation if it is to be secured and maintained. It is in the nature of political authorities that they impose significant costs, and often great sacrifices, on their subjects.

therefore, is the justification he offers for state coercion, not the account of authority itself.

²⁷ Godwin, *supra* note 15, 120.

Not only are members of groups under duties to forbear from certain harmful actions and to come to the aid of each other, but they owe considerable duties to the state itself. They range from the mildly annoying, such as the duties to register one's vehicle and to sit on a jury, to the onerous, such as the requirements to pay one's taxes and to support one's family, as well as the life-threatening, such as the obligations to testify under subpoena against the defendant in a murder trial and to respond to military conscription.

Authorities use many strategies in order to influence people to pay these costs. Perhaps the most cost-effective approach is ideological. By claiming the power to obligate, rather than simply the permission to coerce, authorities attempt to persuade their subjects to comply out of a sense of moral duty. The greater the number of people who accept the law's authority to impose such duties, the fewer the resources that the state must devote to the enforcement of its laws. Once the state has won its peoples' hearts and minds, their bodies will follow. To be sure, legitimation strategies presuppose the existence of considerable machinery of education, indoctrination and manipulation. The dissemination of ideology is, like anything else, not cost-free. Yet, the use of brute force also presumes a vast apparatus of prevention, detection and retribution. Even at the current state of technology, law enforcement is still extremely labor-intensive, requiring a large network of police, investigators, prosecutors and corrections officers, in addition to an extensive system of jails and other correctional facilities. A society that invested all of its resources in intimidation and none in persuasion would quickly bankrupt itself.

But expense is not the only consideration that obliges authorities to legitimate themselves. The machinery that the state erects to reward those who comply and punish those who don't must be manned by at least some who recognize the state's legitimacy and remain loyal to its ideology. It is simply not

possible to have “threats all the way down.” At some point, someone has to carry out the will of the authority because he feels not only obliged, but obligated to do so. The paradoxes of authority can then be turned on those who support this state bureaucracy: why would it be rational for any bureaucrat to heed the directives of authorities to enforce their threats against another? Either the balance of reasons supports enforcing the threat or it doesn’t. If the former is the case, then enforcement is reasonable, but not because the bureaucrat was directed to enforce the threat, but because the balance of reasons supports enforcement; if the latter is the case, then enforcement is unreasonable.

It would seem, therefore, that Ladenson’s approach to authority cannot be defended by the appealing to the fact that, if true, it would supply a solution to the paradoxes of authority. For, as we have just seen, Ladenson’s approach is itself vulnerable to similar problems. Rather than solving the paradoxes, his account suppresses them in one context and fails to prevent them from reappearing in others. The day of reckoning is not averted, just merely postponed.

There are two lessons that we should take away from this discussion. First, it is not possible to solve the paradoxes of authority by attempting to interpret obedience as being mere responses to threats. The enforcement of threats issued by those in power is itself an activity whose rationality and morality must be established. Second, it is important to distinguish between the concepts of having “authority over a person” and of having the “authority to act” or, to put the matter slightly differently, between having authority and merely being authorized. To have authority is to have a normative power to change another’s normative relations. To be authorized is simply to be permitted by someone who has authority to act in a certain manner. The secretary is authorized to open her boss’ mail. But she has no power in this regard. As Raz has aptly put the point, Ladenson’s approach equates what the state may do with

what the secretary may do.²⁸ However, as we have seen, non-normative power over a person and the authority to exercise that power does not add up to, nor can it substitute for, the authority over a person.

B. Theoretical Authorities

According to first horn of the paradox of authority and rationality, if the balance of reasons supports the content of a directive, then the directive is redundant. In response, one might call into question the idea that a directive is irrelevant just because it gives the correct solution to a normative problem. An authoritative directive might be relevant if an agent could use it to solve the problem for herself. By applying the directive, she would be able to arrive at the right conclusion rather than having to deliberate about the merits of the case at hand.

This response attempts to locate the legitimacy of authority in its epistemic instrumentality. Agents have reasons to obey authorities whenever their directives are conclusive reasons to believe that obedience is reasonable. Authoritative directives, in turn, would achieve this epistemic status whenever their source was an expert in the regulated area. Authorities would be legitimate, on this view, just in case their pronouncements constituted expert advice.

This suggests that practical authority, that is, authority concerning what ought to be the case, might be grounded in theoretical authority, that is, authority concerning what is the case. This idea is as old as western philosophy itself. In The Republic, Plato argued that a just society must be governed by philosophers. Plato's view was based both on his high esteem of philosophers and his low opinion of everyone else. The philosophers must rule the just city because only they had access to the truth and

could be trusted to act for the common good. “[T]he simple and moderate desires, guided by reason and right judgement and reflection, are to be found in a minority who have the best natural gifts and best education. . . . This feature too you can see in our state, where the desires of the less respectable majority are controlled by the desires and the wisdom of the superior minority.”²⁹

One need not share Plato’s social theory in order to see the virtues in this approach. Reducing practical to theoretical authority is attractive not only from the perspective of theoretical economy, but also because the rationality of relying on theoretical authority seems unproblematic. When one knows more about a subject than another, it makes good sense to defer to the other’s judgment and, by doing so, one will do better than if one relied on one’s own judgment. Moreover, as we mentioned earlier, reliance on theoretical authority is compatible with autonomy. Someone who acts on authoritative advice may do so because he believes that reliance is supported by the balance of content-dependent reasons.

While this strategy has its benefits, it seems vulnerable to a number of serious objections. First, it is an essential feature of expert-advisee relationships that advisees should not act on expert advice when they know them to be wrong. If I regard the weatherman as a theoretical authority on the weather, I should take an umbrella with me when I see it rain even though he might have forecasted clear skies. There is no value in deferring to theoretical authority when one knows them to be wrong – one relies on theoretical authorities because, and only because, one wants to know what is right. By contrast, practical authorities claim the right to obligate even when they are wrong. In Eliezer’s case, the rabbis required Eliezer to defer even though he knew, and they knew, that they were wrong. This did not deter them,

²⁸ Joseph Raz, “Introduction” in *Authority*, *supra* note 5, 4.

²⁹ Plato, *The Republic* 202 (Desmonde Lee trans., 1955). For a contemporary expression of the epistemic account, see Heidi Hurd, “Challenging Authority,” 100 *Yale Law Journal* 1611 (1991).

however, from asserting their authority. It would seem that practical authority cannot be reduced to theoretical authority, insofar as each have different effects on practical reasoning.

Second, the epistemic approach badly misinterprets the claims of political authorities. Political authorities claim that their directives are more than mere reasons to believe – they claim that they are reasons to act. One does not comply with an order simply by believing that the order was justified; one must act on this belief. Conversely, authorities generally do not punish failures of belief, only failures of action. Unless authorities punish thought crime, it is not against the law to think that the law on any subject is mistaken. Even Hobbes, one of Western society's greatest fans of authority, thought that the commands of the sovereign only bind in *foro externo*, i.e., in action, not in *foro interno*, i.e., in conscience.

It is unlikely, however, that proponents of the epistemic approach will be moved by these arguments. As to the first objection, they will surely respond that this is not so much an argument against, as a wholesale rejection of, their position. Those who reduce practical to theoretical authority are eager to deny that legitimate authorities bind irrespective of content. A critic makes no headway by pointing to the fact that practical authorities have the power to bind even when wrong, given that this “fact” is precisely what the epistemic approach wants to challenge. The Talmudic story, therefore, cannot be used as a refutation of the epistemic position when it has yet to be established that Rabbis' assertion of authority was coherent.

Proponents of the epistemic account admit that subjects can be required to act even when they think that the authorities are wrong. For if authorities are legitimate when their judgments are more reliable than their subjects, they are more likely to be right when their subjects think that they are wrong. The

crucial word here is “think”—when subjects “know” that they are wrong, as when a voice from heaven tells them so, they should not defer.

Proponents of the epistemic approach will also argue that the second objection misses the mark because it badly misrepresents their account of authority. Those who attempt to reduce legitimate practical to theoretical authority do not claim that authoritative directives are simply reasons to believe. They agree that political authorities require action. Their claim, rather, is that when authorities are legitimate, their directives are also conclusive reasons to believe that their content is justified. Indeed, their directives are reasons to act because, and only because, they are reasons to believe. The expertise of legitimate authorities is the ground of their power, not the content of it. Subjects should act on authoritative directives whenever they should believe that their directives are justified.

While these responses are effective, they raise additional problems. According to the epistemic account, A has practical authority over B in domain C if and only if A is more expert than B on C topics. However, if expertise is necessary for legitimate authority, it follows that many areas of the law cannot possibly be legitimate. As is well-known, one of the most vital functions authorities serve is to solve coordination problems: they establish rules of the road, standards for weights and measures, deadlines for filing taxes, etc. In a coordination problem, the parties have an interest in working in a concerted fashion, but given that there are more than one acceptable way of doing so, they must figure out which of these strategies will enable them to coordinate behavior. For example, a motorist wants to drive on the same side of the road as all other motorists. However, because both the left and right sides are equally good choices, he will have problems knowing on which side of the street others will drive and, hence, on which side he should drive. Authorities are able to solve coordination problems because they can designate

one of the strategies as the choice for all to follow. In marking one of the combinations as binding on all, everyone's expectations are focused on that combination and the informational problems are overcome. Motorists know that they should drive on the right side of the road because the law has selected this side by imposing a rule requiring it. A motorist should drive on the right side because he knows that others will drive on the right given that they expect him to do likewise.

The ability of authorities to solve coordination problems, therefore, does not stem from any expertise.³⁰ By hypothesis, coordination problems arise because of the multiplicity of acceptable joint strategies. When the law designates the right side of the road as the proper way to drive, it does not do so because the right side is better than the left – the law was needed precisely because the right side was as good as the left. Legal authorities do not act as experts in this regard because there is nothing over which to exercise their expertise. The law's ability to solve coordination problems stems in large part from the fact that their subjects look to them as the solvers of coordination problems.³¹

Epistemic accounts are deficient because they are unable to account for the authority of law in situations where expertise is irrelevant. Worse still, they cannot legitimize legal authority even in cases where it is highly desirable. The reason is simple: the authority of the law, as opposed to the expert, is impersonal. When a legal official claims authority to issue a directive, the source of that authority resides in the office from which the directive will emanate. From the legal point of view, everyone must listen to President Clinton because he is the President, not because he is Clinton.

³⁰ Expertise might, however, be required in order to identify the existence of coordination problems.

³¹ It might be argued that in coordination problems authoritative directives are reasons to believe that others will act on the coordination solution, even though they are not expert advice. As I will argue in Section VI, C., authoritative directives could not solve coordination problems if it were commonly believed that everyone treated them as evidence of other people's behavior and that they were rational.

By contrast, the authority of an expert is entirely personal. If one should listen to an expert, it is because of the expert's superior knowledge or skill. Smith should listen to Doctor Jones, because he is Jones, not because he is a doctor. The personal nature of expertise is reflected in the fact that theoretical, as opposed to legal, authority cannot be transferred, delegated, acquired, inherited or usurped. Only in a metaphorical sense do experts "appoint" their successors.

Because the legitimacy of an official's authority is impersonal, dependent as it is on the legitimacy of the office, the personal qualities of any official can contribute nothing to the legitimacy of the official's authority. The office is made no more legitimate because of the expertise of its current occupant. To be sure, one has a reason to defer to an official when that official is an expert. But the reason to defer to the official is due to her expert, not official, status. If the law's claim to authority is to be vindicated, the existence of the obligation to obey must not depend on the personalities of those demanding obedience.

III. Preempting Reasons

It would seem that the paradoxes of authority cannot be solved by weakening the concept of authority in the ways previously contemplated. If legitimate authorities exist, they have the normative power to obligate their subjects. We will now explore a less obvious, but far more interesting and sophisticated, attempt at solving the paradoxes. Joseph Raz has argued that the culprit is not the assumption that authorities can obligate, or that they can obligate irrespective of their content, but a false picture about the nature of rationality and the impact that authority has on practical reasoning. Once the distinctive contribution that authoritative directives make to the rationality of compliance, the problems of

authority easily dissolve.

A. The Service Conception of Authority

According to Raz, it is an essential feature of the law that it claims legitimate authority to regulate the conduct of its subjects. It may not possess it, but it necessarily must claim it. As we have just seen, this means, at the very least, that the law claims the normative power to obligate. But, Raz argues, it means much more. To claim such authority is to profess that the exercise of normative power is always directed towards serving the well-being of those who are obligated to obey. Even if the governors inevitably act in their own best interests, they always claim that their job is to serve the governed.³²

In what does this service consist? According to Raz, it is the distinctive task of political authorities to mediate between people and the reasons that apply to them. That is, their role is to consider the reasons that apply to their subjects and to formulate or ratify directives that will enable their subjects to conform to the balance of those reasons. Authorities are legitimate when their subjects are in need of such mediating services and authorities satisfy these needs. Governments exercise their authority, therefore, not by building bridges, educating children or repelling foreign invaders, but rather by producing and validating norms that allow their subjects to conform to Reason.

More precisely, Raz's "service" conception is constituted by two theses about the nature and role of authority. The first concerns the type of reasons that should guide authorities when regulating conduct. According to Raz, all authoritative directives should be based, in the main, on reasons that independently

³² Joseph Raz, The Morality of Freedom, 56 (1986).

apply to the subjects of the directives. Raz call these reasons “dependent reasons” and the above condition the “Dependence Thesis.”³³

Raz distinguishes the Dependence Thesis from the idea that authorities should serve the interests of their subjects.³⁴ A military commander acts on the basis of reasons that apply to the soldiers (i.e., defense of country) even though his commands are not in the interests of the soldiers. Raz also points out that the Dependence Thesis specifies the way that authorities should legislate and adjudicate, not the way they actually do.³⁵ Obviously, many exercises of authority are not based on dependent reasons.

The second half of the service conception concerns the type of argument that must be offered in order to justify authority. According to the “Normal Justification Thesis,” authorities are legitimate when their subjects are more likely to conform to balance of reasons that apply to them if they comply with their directives than if they attempted to conform to that balance directly.³⁶ The legitimacy of authority, therefore, is based on the instrumental rationality of the law. By following the law, agents are more likely do to what they are supposed to do than if they tried to act reasonably by themselves.

The Normal Justification Thesis specifies the condition that authorities must satisfy in order to be deemed legitimate. This is a stringent test and one that probably no authority has ever completely passed.

It is unlikely that in every case some authority will be in a better position than every one of their subjects either to assess the demands of reason or to provide guidance in satisfying such demands. Raz does, however, countenance the possibility of partial legitimacy, i.e., where an authority is legitimate with respect

³³ *Id.* at 42-53.

³⁴ *Id.* at 48.

³⁵ *Id.* at 47.

³⁶ *Id.* at 53.

to some areas of regulation but not others.³⁷ Some authority might, for example, be legitimate with respect to the regulation of worker safety, but not sexual morality. Raz also allows that an authority might be legitimate for some subjects and not for others. I might be bound by certain legal regulations designed to improve physical health, while my doctor might not. Everything depends on whether the subjects will do better by complying with the law than if they tried to conform to Reason directly.

B. The Preemptive Thesis

Perhaps Raz's best-known claim about the nature of authority is that authoritative directives are unlike ordinary reasons in that they are not only reasons to act in accordance with their content, but also reasons to preempt other reasons for action. They are examples of what he calls "preemptive" reasons. When authorities require performance of an action, their directives are not merely added to the balance of reasons, but they also exclude these reasons and take their place. Raz calls this idea the "Preemptive Thesis."³⁸

According to Raz, the Preemptive Thesis follows from both the Dependence and the Normal Justification theses. The Dependence Thesis states that authoritative directives should be based on the balance of dependent reasons. Because these directives are meant to reflect dependent reasons, they cannot be counted along their side. To do so would amount to counting some of the dependent reasons twice.³⁹ Authoritative directives must replace dependent reasons because it is their function to represent them.

³⁷ *Id.* at 73-75

Likewise, the Normal Justification thesis states that authoritative directives are binding just in case subjects would be more likely to conform with the balance of dependent reasons by complying with the directives than if they attempted to conform to the balance directly. If authoritative directives of a legitimate authority did not preempt the underlying dependent reasons, then it would be rationally acceptable for a subject to consider such reasons when deciding how to act. In so doing, however, the subject would be deliberating about the merits of the case and forming a judgment about what ought to be done. Yet, if the authority is legitimate, it is rationally unacceptable for the subject to rely on their own judgements rather than the authoritative directives themselves.⁴⁰ Hence, if the Normal Justification is true, so is the Preemptive Thesis.

Raz illustrates these points using the example of arbitration.⁴¹ Since the arbitrator's decision is supposed to reflect the merits of the case at hand, one who considered it along with the merits would be making a mistake in normative arithmetic, for he would be counting some of the merits twice – once directly and once through the decision that incorporates their force. Moreover, parties use arbitrators to resolve their disputes whenever they find that they are unable to resolve their conflict by themselves, or when doing so would be too costly. They give the arbitrator the power to decide authoritatively who is right and, in so doing, they relinquish the right to challenge the decision. To consider the arbitrator's decision as an ordinary reason to be added to the balance along with the other merits would vitiate the purpose of the arbitration. The decision is supposed to eliminate the need to deliberate and debate the merits of the case that they have submitted.

³⁸ *Id.* at 57.

³⁹ *Id.* at 58.

⁴⁰ *Id.* at 61.

Because authoritative directives preempt only those reasons they are meant to reflect, they will lack preemptive force in situations where authorities are unable or unwilling to track the balance of dependent reasons. An arbitrator's rulings, for example, need not be followed if she is drunk or has been bribed or if new evidence of great importance later pops up.⁴² Likewise, because authoritative directives have preemptive force in order to prevent agents from acting directly on the merits, directives will fail to preempt just in case subjects can determine that an error is made without scrutinizing the underlying merits of the case. Raz argues that authoritative directives may not be binding if they are "clearly wrong." A clear mistake is a mistake that may not deviate substantially from the balance of reasons but wears its error on its face. "Establishing that something is clearly wrong does not require going through the underlying reasoning."⁴³ This is to be distinguished from a great mistake which does indeed deviates greatly from the balance of reasons and whose detection as an error requires the agent to deliberate on the underlying dependent reasons putatively supporting the claim.

Raz illustrate the difference between the two types of mistakes using the case of adding integers. If an authority tells the agent that the sum is an integer, the only way to detect a great error in the sum may be to actually add all of the integers and then compare the results. On the other hand, if the answer presented were a fraction then it would be clear that the authority had made a mistake. In such a case, the dictates of rationality do not require that the agent believe clear mistakes, but would urge the acceptance of great mistakes.

⁴¹ *Id.* at 41-42.

⁴² *Id.* at 42.

⁴³ *Id.* at 62.

C. The Justification of the State

Raz mentions five scenarios where political authorities are able to achieve legitimacy under the Normal Justification Thesis.⁴⁴ The first involves cases where expert advice is needed and authorities possess the requisite competence. Regulations that deal with the approval of pharmaceuticals, for example, are typical based on the expertise and information that the government possesses but that the ordinary citizen lacks. By deferring to such directives, subjects are likely to do better than if they tried to figure out for themselves which drugs are safe and efficacious.

The next two classes of cases involve situations where authorities compensate for various shortfalls of rationality. In some situations, authorities may be less easily swayed by temptation and bias than their subjects. In other cases, deliberation may be costly, either because it generates anxiety, fatigue or redirects cognitive and emotional resources from other endeavors. By relying on authorities, subjects will be able to avoid the costs that surround the attempt to conform to the balance of the reasons that apply to them.

The fourth class concerns cases where authorities are in a superior position to provide solutions to coordination problems. Authorities, for example, are generally better at setting the rules of the road than are drivers. Although not always true – sometimes informal conventions are more efficient – it is often the case that top-down solutions are superior to bottom-up ones, and when they are, authoritative solutions will be legitimate and binding on participants.

The last class concerns prisoner's dilemmas. In a prisoner's dilemma, each player does better if

all cooperate than if all defect; in contrast to coordination problems, however, each does best if the others cooperate, but he defects. Therefore, if everyone acted so as to achieve the best result, i.e., unilateral defection, they will bring about an outcome that is individually suboptimal, i.e., universal defection. In order to achieve the efficient solution, authorities can issue directives that require the “players” to “cooperate.” So, while it might be rational for me not to pay my taxes if there were no law requiring me to, and hence no one would pay their taxes, I will have a reason to pay my taxes when there is a law to requiring me to, and hence enough people will pay their taxes.

D. Razian Solutions to the Paradoxes

An enormous virtue of Raz’s theory of authority is that it provides powerful solutions to the paradoxes of authority. The solutions, as one might guess, rest on the two distinctive claims that Raz makes about authority: first, that they are justified primarily on instrumental grounds, and second, that their directives have preemptive force.

Before we discuss Raz’ solutions to the paradoxes, however, we must first translate his terminology into ours. I will assume that when Raz speaks of “dependent” reasons, he is referring to what we have called “content-dependent” reasons. Aside from linguistic affinities, both sorts of reasons share the same feature: they are reason that apply to subjects independently of the existence of authoritative directives.

With respect to preemptive reasons, however, Raz denies that they are the same as peremptory

⁴⁴ *Id.* at 75.

reasons. A preemptory reason is a reason not to deliberate about other reasons, whereas a preemptive reason is a reason not to act on other reasons. Raz faults Hart for thinking that submission to authority requires the actual surrender of judgment, instead of just the forfeit of the right to act on that judgment.⁴⁵ One obeys an order even when one thinks that directive is wrong.

I think that Raz is misinterpreting Hart's, indeed even the ordinary, notion of "deliberation." To deliberate is not simply to engage in the thought process of weighing pros and cons. Rather, deliberation is essentially action-guiding. One who deliberates does so with the aim of forming an intention to act on the results of that deliberation. A preemptory reason is not, then, a reason not to think about other reasons, but a reason not to form an intention based on them.

In another way, though, Raz is right that a preemptive reason is different from a preemptory one. A preemptive reason is not simply a reason not to act on other reasons. It is a reason that replaces those reasons. A preemptory reason, by contrast, simply excludes certain reasons from serious consideration. It does not replace them in deliberation.

Nevertheless, a preemptory reason that is also content-independent, i.e., a CIP reason, has the same normative force as a preemptive reason. For a preemptory reason will exclude only those reasons that are content-dependent and thus will not, so to speak, exclude itself. A CIP reason will therefore be a reason to exclude other conflicting reasons as well as a reason for acting in accordance with its content, just like a preemptive reason.

We are now in a position to discuss the paradox of authority and rationality. Assume that a legitimate authority issues a directive to some agent requiring him to do an action that is already supported

by the balance of reasons. The agent should surely conform to the directive in this instance. But is its existence relevant to his practical reasoning? It is on Raz's theory. According to the Normal Justification Thesis, since the authority is legitimate, its directives make it more likely that the agent will do better in terms of the reasons that apply to him by complying with the directives than if he tried to act in accordance with these reasons directly. The directive, therefore, is a reason for action, because through its guidance, the agent increases his chances of acting in accordance with the balance of content-dependent reasons.

Assume, on the other hand, that the content of the directive is not supported by the balance of reasons. Should an agent follow the directive? The answer here is "Yes" again, for according to the Preemptive Thesis, directives issued by legitimate authorities are not merely added to the balance of dependent reasons but replace these reasons. As a result of this exclusion, the dependent reasons that counsel against conforming to the directive are no longer relevant. The only reasons that count are those that are left in the balance, which, in this case, happen to be the authoritative directive.

The possibility of rational obedience to legitimate authority is thereby secured. If an agent believes that he will do better in terms of the reasons that apply to him by deferring to directives issued than by deliberating, he is rationally required to defer to each directive irrespective of his judgments about the balance of content-dependent reasons.

Unfortunately, when we turn to the paradox of authority and autonomy, we must resort to speculation. The reason is that Raz's understanding of autonomy differs from the one we have been considering. Raz sometimes considers autonomy to be a principle of practical reason, other times to be a capacity for, or the exercise of, self-determination. When understood as practical principle, it is a rational,

⁴⁵ *Id.* at 39.

not a moral, one. “[O]ne’s right and duty to act on one’s judgement of what ought to be done, *all things considered* ... I shall call the principle of autonomy.”⁴⁶ Raz adds in a footnote that “[i]t is clear that this principle of autonomy is not really a moral principle but a principle of rationality.”⁴⁷ Under this conception of autonomy, the paradox of authority and autonomy is a simple variant of the paradox of authority and rationality and can be resolved in the same way.

When authority is seen as a capacity for self-determination, it is inevitable that authority and the exercise of autonomy will clash. If one’s actions are self-determined, they cannot, at the same time, be determined by authority. Raz recognizes this conflict, but is not particularly alarmed by it, seeing it as involving a tradeoff. Many times, one should sacrifice one’s right to act on one’s judgments when they are inferior to another’s. Yet, it is sometimes better to act on one’s inferior judgments just because they are one’s judgments. Without making mistakes, one can never develop the capacities necessary for autonomous action in other spheres of life.

Although Raz does not consider autonomy in the same way we have, that is, as a thesis about the space of reasons, I think it is clear that Raz is sympathetic to the same basic concerns. After all, it is core commitment of his service conception that legitimate authorities are not deontically creative: they do not have the ability to create moral duties that their subjects did not have previously. “[A]uthorities do not have the right to impose completely independent duties on people ... their directives should reflect dependent reasons which are binding on those people in any case.”⁴⁸

While we don’t know his response to the paradox of authority and autonomy, we can construct a

⁴⁶ Joseph Raz, The Authority of Law 27 (1979).

⁴⁷ *Id.*

⁴⁸ Authority, *supra* note 5, 135.

“Razian” response. A Razian might deny that an agent is heteronomous just because they act on a preemptive reason. Heteronomy results not from acting on CIP reasons, but from knowingly failing to conform to content-dependent reasons. If a directive passes the test laid out in the Normal Justification Thesis, then its content-independent and peremptory nature is justified by its ability to engender conformity with the balance of content-dependent reasons for action. One can, therefore, act on a CIP and content-dependent reasons at the same time. One would not be obeying simply because another told one to do so, but because by doing as one is told, one would be more likely to be acting on undefeated content-dependent reasons.

IV. Is Preemption Necessary?

Raz’s theory provides a powerful response to the paradoxes of authority, as we have seen, because it combines an instrumental approach to authoritative directives with a hierarchical theory of rationality.⁴⁹ On the one hand, it is the function of authoritative directives to maximize conformity to the balance of content-dependent reasons. Yet, it is not necessary for any authoritative directive to actually fulfill such a function. As long as it is rational for someone to accept an authority as legitimate within a certain domain, it is rational to guide one’s conduct by any directive issued concerning that domain, even when conformity to such a directive is not, by the agent’s own lights, supported by the balance of content-dependent reasons.

⁴⁹ Raz’s theory is notably similar to Rawls’ presentation of rule-utilitarianism. See John Rawls “Two Concepts of Rules” in Theories of Ethics (Phillipa Foot ed., 1967).

The success of Raz's response depends, then, on whether an instrumental approach to authoritative directives is compatible with a hierarchical theory of rationality. The worry here is this: if authoritative directives are supposed to maximize conformity to the balance of content-dependent reasons, how one have a reason to guide one's behavior by a directive when it does not do what it is supposed to do? As the act-utilitarians have argued in a similar context, to follow through on a rule when it gives suboptimal results is a form of "rule-worship."⁵⁰ If authoritative directives are instrumental reasons for action, they cannot, at the same time, be preemptive reasons for action; alternatively, if the Normal Justification Thesis is true, then the Preemptive Thesis cannot be.

As we have seen, Raz presented two arguments to show that the instrumentality of directives entails its preemptive effect—one based on the Dependence Thesis and the other on the Normal Justification Thesis. Critics have responded primarily to the second of these arguments. They have attempted to show that authoritative directives can be instrumentally valuable even if they do not preempt the reasons they are supposed to reflect. In this next section, we will explore this possibility.

A. The Simple Model

To recall, Raz argued that the Preemptive Thesis follows from the Normal Justification Thesis: if authoritative directives did not preempt dependent reasons, it would be rationally acceptable for a subject to try to conform to the balance of dependent reasons, which is precisely what the Normal Justification

⁵⁰ See, e.g., J.J.C. Smart, "Outline of a Theory of Utilitarian Ethics", in Utilitarianism: For and Against (J.J.C. Smart and Bernard Williams eds., 1973).

Thesis declares to be rationally unacceptable. Notice that this argument assumes that there are only two options – either subjects completely defer to an authoritative directive or they completely ignore it and deliberate in its absence. Yet, these two alternatives do not seem to exhaust all the possibilities.

Consider, for example, cases where practical authorities are theoretical authorities on a certain subject. Authoritative pronouncements are, then, reasons to believe that the balance of content-dependent reasons supports conformity. To comply with an authoritative directive might, then, involve treating it as strong evidence regarding the balance of content-dependent reasons and to act on the basis of it, as well as all of the other available evidence. Authoritative directives would not preempt the reasons they are meant to reflect – they would be additional reasons that lend their support to the pro-content side of the balance and would be considered along side all of the other content-dependent reasons.

Such an approach is perhaps the simplest of the many models mentioned in the literature. Let us call it the “Simple Model.” The Simple Model treats authoritative directives as it would any other reason: as a first-order normative consideration that is added to the balance of reasons and that can be considered with all other content-dependent reasons. In the case of legitimate authority, the weight of the authoritative reasons is great – normally great enough to outweigh any other contrary reason.

At least at first glance, the Simple Model easily handles situations where practical authorities are also theoretical authorities: it treats their directives as strong first-order reasons to believe that the balance of content-dependent reasons supports conformity. A similar analysis can be given for directives that are used to cut down on deliberation costs and to compensate for cognitive incapacities. In these cases as well the directives are weighty first-order reasons to believe that the balance of reasons supports the content of the directives and are normally strong enough, given the costs of deliberation on the merits, to

outweigh any contrary reason.

With regard to coordination problems, the Simple Model sees authoritative directives as creating salience. The equilibrium specified by the content of the directive is made more conspicuous by the issuance of the directive and this “marking off” focuses the attention of all players on that solution. Each player not only focuses on that solution, but expects that others will similarly focus and expects that others will expect them to similarly focus. As with the other cases, authoritative directives do not preempt the underlying reasons supporting one solution over another – they are simply very powerful reasons to act accordingly which normally outweigh the reasons to act differently.

Advocates of the Simple Model argue that, contrary to Raz’s argument, one who treats authoritative directives as the Simple Model suggests would not undo the benefits of relying on legitimate authority. For, read through the lens of the Simple Model, the Normal Justification Thesis states that an authority is legitimate just in case assigning significant weight to their directives enables one to do better than if one did not assign it any weight at all. An authority is justified in coordination cases, for example, just in case treating the authoritative directive as marking off the salient strategy and hence adding it to the balance of reasons allows one to do better than if one did not treat the content of the directive as salient but tried to solve the coordination problem by oneself. The Normal Justification Thesis would not, therefore, entail the Preemptive Thesis, given that the instrumentality of authority can be had even without preemption.

B. Variations: Transformation, Reweightings, and Presumptions

The Simple Model is but one of many accounts offered as alternatives to Raz's Preemption Model. Some proposals represent subtle variations on the Simple Model, while others propose more radical departures that closely resemble Raz's approach. What they all have in common is that they reject the Preemptive Thesis, i.e., that authoritative directives must completely preempt the reasons they are meant to reflect.

Beginning with subtle variations of the Simple Model, some theorists have argued that authoritative directives have transformative powers: it is their function to alter the set of content-dependent reasons. When, for example, authoritative directives constitute strong evidence that the balance of content-dependent reasons supports conformity, they are weighty reasons to believe that other reasons are not in fact reasons and hence are not entitled to their place in the balance. On this Transformative Model, authoritative directives that are used to compensate for irrationality do not preempt any content-dependent reason. They are reasons not to act on untrustworthy beliefs that there are reasons to act contrary to the advice, not as reasons to remove a *bona fide* reason from the balance of reasons. As Heidi Hurd argues, "[o]ne's condition of incompetence is but evidence of the fact that in working out these content-dependent reasons, one may not be employing true premises."⁵¹

In coordination cases, the authoritative directive focuses attention away from other options to the content of the directive. Whatever reasons players had for acting on the other options, they no longer

⁵¹ Hurd, "Challenging Authority," *supra* note 29, 1624. See also D. S. Clarke, Jr., *Exclusionary Reasons*, 62 Mind 253 (1977).

have. Consider Leslie Green's analysis of coordination norms:

Given a generally shared expectation that one alternative will be followed, there is no longer any appeal whatsoever in [acting in different way], for in doing so one would be swimming against the tide which, by hypothesis, one has no interest in doing. But note: these non-options (i.e., those that are not salient) leave no practical trace--one does not hanker after them, and they exert no residual attraction from any point of view; they are simply outweighed. To achieve an equilibrium by appealing to or creating a conventional norm, one need only act on the balance of first order reasons.⁵²

More complicated are proposals that treat authoritative directives as second-order reasons that affect the weight of the first-order reasons that they are meant to reflect. Stephen Perry calls such reasons "reweighting" reasons.⁵³ A reweighting reason is a reason to act "as if" another reason had a certain weight. A subject who regarded an authoritative directives as strong evidence as to which actions were reasonable might, according to Perry, reweight the content-dependent reasons, transferring some of the weight from reasons that don't support the content of the directive to others that do.

Perry's reweighting reasons are generalizations of Raz's preemptive reasons. An exclusionary reason is a limit case of a reweighting reason, one that transfers all of the weight of certain reasons that do not support its content to those that do. Aside from preemptive reasons, therefore, reweighting reasons do not preempt the reasons upon which they operate.

In addition to functioning as reweighting reasons, Perry also claims that authoritative directives

⁵² Leslie Green, The Authority of the State 113-114 (1988). For similar analyses, see Donald Regan, "Authority and Value: Reflections on Raz's Morality of Freedom," 62 Southern California Law Review 995, 1027 (1989); Heidi Hurd, "Sovereignty in Silence," 99 Yale Law Journal 1016-19 (1990); Larry Alexander, "Law and Exclusionary Reasons," 18 Philosophical Topics 7 (1990).

⁵³ Stephen Perry, "Second Order Reasons, Uncertainty and Legal Theory," 62 Southern California Law Review 913 (1989). See also, Stephen Perry, "Judicial Obligation, Precedent and the Common Law," 7 Oxford Journal of Legal Studies 215 (1987).

may take the form of presumptions, or as he calls them, “epistemically-bounded” reasons.⁵⁴ On this model, those who rely on authoritative directives presume that the content of the directives are supported by the balance of reasons. Such presumptions do not completely preempt the underlying reasons, insofar as they can be rebutted in certain circumstances. Perry believes that authoritative directives establish thresholds of credibility that preempt inquiry into the underlying reasons just in case no reason of a sufficiently weighty sort exceeds the threshold. Other advocates of the Presumption Model, such as Fred Schauer, believe that these presumptions can be rebutted by taking a “perfunctory glimpse”⁵⁵ at the content-dependent reasons to determine whether there exists a good enough reason to doubt the reliability of the presumption.

As with the Simple Model, these alternative proposals permit subjects to benefit instrumentally from authority without the need for preemption. Suitably reinterpreted according to these models, the Normal Justification Thesis legitimates authority just in case treating authoritative directives as weighty first-order/second-order reweighting/presumptive reasons and then acting on the basis of the resulting balance of reasons the subject is more likely to conform to the balance of content-dependent reasons than if he did not treat them in this manner. In situations where such a condition is satisfied, a subject who acts on the basis of her judgments about the balance of reasons would not be undoing the work of authority, but rather would be harnessing the value of authority and using it as it ought to be used.

⁵⁴ See Perry, “Second-Order Reason,” *supra* note 53, ??.

⁵⁵ See Fred Schauer, Playing by the Rules: An Examination of Rule-Based Decision-Making in Law and in Life 91 (1991).

C. Is Preemption Rational?

Not only have Raz's critics argued that the instrumental benefits of authoritative directives can be secured without preemption, they claim that people do not generally treat authoritative directives as preemptive reasons.⁵⁶ It is a well-known fact about certain legal systems, for example, that judges have the power to depart from established rules. Common Law judges may refuse to follow a precedent under certain circumstances, e.g., if the rule is obsolete or sufficiently unjust. But if judges treat legal rules as preemptive reasons, and if preemptive reasons always defeat reasons which fall within their scope no matter how strong the first-order reasons, a judge would be barred forever from acting on reasons that favor departing from the rules. It would then seem as if the preemptive nature of legal rules is incompatible with the revisibility of the Common Law.

Not only do judges not treat past precedents as preemptive reasons, but Raz's critics argue that they should not. Common Law doctrines would become entrenched and the flexibility of the law to adapt to new circumstances would be greatly diminished. Michael Moore, for example, claims that "Raz' account suffers from [this] defect ... some cases that a judge ought to overrule won't be overruled. ... The judicial obligation is to overrule whenever the balance of reasons (including the reasons inclining against overruling given by the rule-of-law values) make it the right thing to do."⁵⁷

By contrast, on the alternative models we have been discussing, judges should depart from legal rules just in case reasons existed that were powerful enough to override the reasons to stick to such rules.

⁵⁶ See, e.g., Stephen Perry, "Judicial Obligation," *supra* note 53; Schauer, *supra* note 55, 91.

⁵⁷ Michael Moore, "Authority, Law and Razian Reasons," 62 Southern California Law Review 827, 867 (1989). See also, Perry, *supra* note 53, 963.

Similarly, subjects should disobey any directives that were not supported either by the balance of first-order reasons, the balance of reweighted first-order reasons or the presumed balance of first-order reasons, respectively. The benefit of these accounts, therefore, is that they attempt to account for virtues of relying on authority without succumbing to the vice of rule-worship.

We might sum up these criticisms by seeing them as attacking Raz's solution to the paradox of authority and rationality. When the balance of content-dependent reasons supports conformity to an authoritative directive, Raz has not shown why authoritative directives must be understood as preemptive reasons rather than weighty first-order reasons, reweighting second-order reasons or presumptions. On the other hand, when the balance of content-dependent reasons supports non-conformity, Raz has not shown how authoritative directives can make normative considerations that are otherwise relevant in their absence completely irrelevant in their presence.

D. Double Counting and Prisoner's Dilemmas

Raz's critics have largely ignored his argument from double-counting. They have also failed to respond to his analysis of authoritative solutions to prisoner's dilemmas. In this section, I would briefly like to examine whether these arguments are effective.

According to the argument from double-counting, authoritative directives must be preemptive in order to avoid counting dependent reasons twice. Because authoritative directives are supposed to reflect dependent reasons, these reasons cannot have independent weight along with directives in the balance. By attempting to reflect dependent reasons, they must replace them.

This argument, however, is both too weak and too strong. It is too weak in that it establishes only that when a subject is certain that the directive tracks the balance of dependent reasons, he should not count the directive and the dependent reasons together. This indeed would involve double-counting. However, if a subject is not convinced that the directives issued will fully reflect the balance of dependent reasons, then an agent might consider the dependent reasons as a check against the possibility of error. A subject might reason as follows: “Because the authority in question is highly reliable, I will give great weight to the authoritative directive in my deliberations. But since there is a chance that the authority might have made a mistake, I will only also consider any reasons that might militate against obedience when judging how to act. If there is a reason sufficiently great to disobey, then I will conclude that the directive was not successful in reflecting the balance of reasons and I will disobey.” The dependent reasons are not here being counted twice – rather they being used to ensure that they are at least counted once.

The argument is also too strong, for if valid, all forms of advice would have preemptive status. Those who dispense ordinary advice also try to base their recommendations solely on dependent reasons. It would follow, according to the double-counting argument, that ordinary advice is also a preemptive reason, for it cannot be considered a first-order reason for belief without counting reasons twice. This would lead to absurdity – any piece of credible advice would defeat all of the advisee's reasons for belief as long as he knew that the advisor considered them in her judgment.

As for Raz's analysis of prisoner's dilemmas, recall that he argued that authoritative directives can be used to solve such problems if they preempted the reasons for defecting. While he believes that preemption by an authoritative directive is sufficient to solve a prisoner's dilemma, he clearly does not think that they are necessary. As Hobbes argued, authorities can solve prisoner's dilemmas by

sanctioning defection, i.e., by realigning preference so that defection is not only socially, but individually, suboptimal. What is unique about Raz's analysis is his suggestion that sanctions are not the only way to bring individual and social rationality in line – authoritative directives may also accomplish through their preemptive power.

While this is an intriguing proposal, much more needs to be said in order to overcome the following objection. In a prisoner's dilemma, it is in each player's interest to defect instead of cooperate. If Reason requires that agents act on their interests, then it would seem that an authoritative directive demanding cooperation would be demanding action contrary to Reason. However, if Reason instead requires agents to act for the collective good – either because acting on the collective good is a good itself or is an indirect way of maximizing conformity to one's own interests – then players ought to cooperate regardless of whether there is an authoritative directive.⁵⁸ Guidance by authoritative directives once again seems to be either irrational or irrelevant.

V. The Decision and Constraint Models

The debate between Raz and his critics is a long-standing one and it is not possible to do justice to its subtlety and complexity in this essay. Rather than rehearse the responses and counter-responses in that dialectic, I would like to present an argument that, to my knowledge, has not been previously made and that aims to show that, at least in certain circumstances, authoritative directives cannot be preemptive

⁵⁸ To be sure, authoritative directives may help coordinate behavior between players so that each can act on the same cooperative solution. Directives might even make the players aware that they are involved in a prisoner's dilemma. However, authorities would be not be solving prisoner's dilemmas *per se*, but rather coordination problems or

reasons. If authoritative directives are capable of serving their functions, it is not in virtue of their alleged preemptive force.

This argument, however, will give no solace to Raz's critics. For I will claim that this argument demonstrates that the models advanced by Raz's critics are inadequate as well. The mistake made by all of these accounts of authority is their assumption that willing obedience to authoritative directives is a form of decision-making.

As an alternative account, I will suggest that authoritative directives are instrumentally valuable when, and only when, they are capable of affecting the feasibility of non-conformity. When a subject commits himself to following an authority in order to benefit from its directives, that person constrains his future self to act on the demands of the authority, whatever they may be. The subject does not choose to obey when the directive is issued – when ordered, the subject has no choice but to obey.

In the following sections, I will generalize the suggestion to other cases of authoritative guidance. I will argue that authoritative directives are instrumentally valuable to rational agents if and only if they are causal constraints on action. I will try to show, in other words, that authoritative directives could not perform the functions they are thought to serve if subjects had, or believed they had, the choice whether or not to obey. I will then argue that what I call the Constraint Model is the key to solving the paradoxes within a framework of authority that ties legitimacy to the instrumentality of authoritative directives.

A. Is Preemption Sufficient?

Charlie wakes up one morning, notices that his roommate Larry is overweight and gives him some friendly advice. “Hey, Larry, you’re pretty fat. You really need to work out.” Larry looks over at Charlie and repays the compliment. “Charlie, you ain’t looking too good yourself. It wouldn’t hurt you to visit the gym once in a while either.” Unfortunately, each knows that the other speaks the truth. They both realize that they should lose weight and that the only way to do this is by going to the gym and exercising.

Turning over a new leaf, Charlie decides to stop by the gym on his way to work every day and exercise. Larry, however, is worried. He knows that he has tried to follow such a regimen in the past but has always failed. After voicing his concerns to Charlie, Charlie suggests that Larry hire a personal trainer. Larry is encouraged by the thought and so hires Sonnie, an ex-marine drill sergeant, for the necessary motivation.

The next morning, promptly at six o’clock, Sonny arrives at the apartment to pick Larry up for the gym. Larry says that he is very tired and, although he knows that he should work out, he tells Sonny to come back tomorrow. Sonny barks back: “If you don’t commit yourself to following my orders, you will never get into shape. And I won’t be involved with a loser.” Larry sees the wisdom in Sonny’s plan and so commits to following his every command.

Assume that Sonny orders Larry to go to the gym. According to Raz’s theory, Larry will regard this directive as a reason for not acting on the reasons for staying home. But can he have such a reason?

An authoritative directive is a preemptive reason if and only if the subject is likely to do better by preempting certain reasons and acting on the resulting balance than if he tried to comply with the first-order reasons directly. However, regardless of whether Larry preempts certain reasons or not, the balance points in the same direction, i.e., towards the option of going to the gym. After all, the reasons for staying home are outweighed by the reasons for going to the gym and Larry is painfully aware of this. Indeed, that is why Larry hired Sonny – because he knows that he should go to the gym to lose weight instead of lazing around the house. Doesn't this show that the order isn't a reason for not acting for certain reasons, because even if Larry considered these first-order reasons in deliberation, Reason would still recommend that he work out?

This same point might be made by comparing Larry's situation with that of Charlie's. Both have the same first-order reasons for action. Both have reasons to go to the gym that are stronger than those for not going and both know this. From the perspective of the first-order balance of reasons, there is no difference between the two. Yet, only Larry needs Sonny. Hence, the value of authority to Larry cannot be traced to the benefits of preemption given that preemption would give the same results for Charlie as well.

Larry's case is not simply a problem for Raz's theory – it generalizes to all other accounts that we have surveyed. According to the Simple Model, for example, Sonny's order is a first-order reason for Larry to comply if and only if Larry would likely do better by assigning the order significant weight and acting on the resulting balance than if he did not assign it such weight. If so, then Sonny's order is not a reason for action because the balance with or without the directive points towards going to the gym.

Why do all of the models we have discussed fail to account for the normativity of Sonny's order?

I believe that the problem stems from several tacit assumptions made by all of these models. The first assumption is that people who submit themselves to authority are free not to follow them if they so wish. Nevertheless, these people choose to obey. Each act of compliance involves a choice to comply.

The second assumption is that authoritative directives affect practical reasoning by affecting a subject's preferences over options, or beliefs about those preferences. Therefore, when an agent accepts an authority's legitimacy and is ordered to act in a certain manner, the agent must believe that he prefers conformity to non-conformity. Moreover, if he were not so ordered, he might have come to the opposite judgement.

Putting these two assumptions together we get: when a subject obeys a directive issued by an authority deemed legitimate, he chooses to obey because he now prefers to conform than not conform. For example, when someone heeds a command issued by someone deemed to be a theoretical authority, that person uses the pronouncement as some sort of reason to believe that he prefers conformity to non-conformity and decides to act on that belief. Let us call this account of authoritative guidance the "Decision Model."

The Decision Model has great intuitive appeal. Unfortunately, it is also false, as Larry's case demonstrates. If the function of authoritative directives were to affect preferences, or beliefs about preferences, then Sonny's order would be pointless. By hypothesis, Larry prefers and knows that he prefers to go to the gym. He does not need to rerank his options – they are in perfect alignment. Larry is not having trouble with his preferences but with his ability to act on those preferences. The Decision Model is unable to explain how authorities can help people like Larry.⁵⁹

⁵⁹ As Sidney Morgenbesser pointed out to me, people like Larry who know what they ought to do, but are unable to do it, are known in Yiddish as "schlemeals."

B. Constraining Your Future Self

I would like to suggest that the Decision Model cannot account for the normativity of Sonny's order because it ignores the essential "volitional" aspect of directives. Directives are not tools for making decisions – they are a way of preventing decisions from being made. When someone submits to authority, the aim is to constrain his future self to act in accordance with the demands of some third party. Should that act of commitment be successful, the agent becomes unable to act contrary to will of the authority.

The suggestion is that submitting to authority involves trying to do to yourself internally what Ulysses was able to do externally when he lashed himself to the mast. It is to forgo later choice by the operation of the Will, but it is as real as using some precommitment mechanism.⁶⁰ According to what I shall call the “Constraint Model,” authoritative pronouncements are relevant to practical reasoning because, and only because, they affect feasibility. A directive, once issued, is not a factor to be considered in future deliberation about whether to comply. After submission, the agent no longer deliberates about whether to comply. The subject merely figures out what counts as implementation of the rule, not which action will maximize expected utility.

Understood game-theoretically, a rational agent who must consider his future actions is strategically interacting with another agent – his later self. When an agent submits to authority, his present

⁶⁰ See, e.g., Stanley Milgram, Obedience to Authority, 134 (1974) (“Since the agentic shift is largely a state of mind some will say that his shift in attitude is not a real alteration in the state of the person. I would argue, however, that these shifts in individuals are precisely equivalent to those major alterations in the logic systems of the automata considered earlier. Of course, we do not have toggle switches emerging from our bodies, and shifts are synaptically effected, but this makes them no less real.”) But see Section VI., D., where I argue that there are some important differences between authoritative submission and pre-commitment.

self attempts to constrain the actions of his future self. In this “game,” the present self makes the first move and the future self will be barred from acting contrary to such a rule if the present self’s actions are successful.

Several clarifications should be made about this proposal. The Constraint Model does not maintain that someone who submits to authority can never disobey. This, of course, would be absurd. When someone submits to an authority, they must sincerely attempt to constrain their future selves. That does not mean that they have constrained themselves, only that they attempted to do so. They might be wrong – the constraint might not have taken. Or it might take, but later lapse. The Constraint Model deals only with successful submission, where the agent actually follows through on the directives issued.

The Constraint Model also does not hold that authoritative directives cause us to conform to them. Because obedience is an intentional action, the subject has to act and act for a reason. Rather, according to the Constraint Model, authoritative directives causally constrain non-conformity – they prevent us from breaking them. When a subject obeys an authority deemed legitimate, she acts for reasons even though she does not make choices by so doing. Obedience is intentional, but not free, action.

Because the Constraint Model locates the functionality of authoritative directives in their ability to affect the feasibility of non-conformity, it is able to account for the instrumental value of Sonny’s order. As we mentioned, Larry’s problem lies not with his preferences but with his ability to act on those preferences. Sonny’s order enables him to act on his preferences by cutting off his ability to act on temptation. By ordering Sonny to the gym, staying home becomes infeasible – working out becomes the only available option. Charlie, on the other hand, has no need for Sonny because the feasibility of staying

home poses him no practical difficulties. His knowledge that this option is suboptimal is sufficient to prevent him from exercising the option.

Not only can the Constraint Model account for the instrumental value of Sonny's order, but it also easily solves the paradox of authority and rationality. To see this, assume that an agent is committed to an authority in order to benefit from its directives. Assume further that the authority issues a directive that, according to the agent's own lights, is not supported by the balance of reasons. Is it rational for the agent to comply? According to the Constraint Model, it is. For, on this view, when an agent submits to authority to benefit from its directives, that agent has no choice but to apply the directive when she recognizes that it is applicable. Compliance is the only feasible option and, hence, is the only optimal option.

Assume now that the directive is, by the agent's own lights, supported by the balance of content-dependent reasons. Is the directive relevant to the agent's practical reasoning? Again, on the Constraint Model, it is. According to the Constraint Model, authoritative directives affect practical reasoning by transforming the set of feasible options, not the preferences over those options, or beliefs about those preferences. Hence, even if the agent preferred to conform prior to the issuance of the directive, once the directive is issued, it will leave its practical mark – what once was feasible is no longer feasible.

The Constraint Model also constitutes an effective response to the paradox of authority and autonomy. The will of the another possesses normative power, on this view, because of its causal power.

The directives of the authority constitute reasons to obey given that the directives render disobedience infeasible – the directives make conformity the best options by default. Ironically, authority is reconciled with autonomy by showing how committed subjects are not autonomous, in the sense of not being in full

control of their actions.

To be sure, it is morally permissible to abdicate control over one's actions only if one has a good moral reason to do so. One cannot absolve oneself of responsibility simply by claiming that one had no choice but to follow orders, when one made the choice not to have a choice. We have seen one example where agents can have good reasons to constrain their future selves to heed the demands of others. In the next sections, we will see others.

VI. The Inadequacies of Decision Models

The Decision Model is incapable of explaining the normativity of Sonny's order. This demonstrates, at the very least, that it cannot provide a general framework for understanding the rationality of authoritative guidance, insofar as it is false in some contexts. However, this might not appear so damaging. One could argue that cases like Larry's are unusual and that authorities are normally unnecessary in situations where agents know what they ought to do. They are valuable, for the most part, in cases where agents require information of some sort, either because they lack expertise, cognitive resources or advantageous positioning to coordinate behavior.

In the next sections, I will argue that not only is the Decision Model false in some cases, it is false in all cases. Authoritative directives cannot furnish expert advice, compensate for shortfalls in rationality, or coordinate behavior if subjects used authoritative directives to make decisions in a rational manner.

A. Authoritative Advice

In order to show that authoritative advice would be useless if advisees were free not to accept the advice, I would like to present an argument of Isaac Levi's. Levi has shown that it is normally not possible to harness the informational value of theoretical authorities by choosing to update one's epistemic state via Bayesian Conditionalization so as to accommodate their reliable evidence.⁶¹

Levi's argument can be made clear by the following example. Suppose Tony is sick and must decide whether to take a certain antibiotic he finds in his medicine chest. Tony has very little medical knowledge; in fact, he is so unsure about what to do that he can't even assign probabilities to the proposition, call it h , that the pill will make him feel better. That is, he does not even assign odds to h , but rather suspends judgment on what the appropriate odds should be. Levi claims that in this situation Tony's probabilities are "maximally indeterminate." His credal state should be represented by the set of all possible probability distributions.

Suppose that Tony believes that his doctor is right in 95% of his medical judgments. It would seem reasonable for Tony to trust his doctor's recommendation's about the antibiotic and then take the pill. However, as Levi points out, if Tony treats his doctor's recommendation as evidence and updates by Bayesian Conditionalization,⁶² he will learn nothing and his probabilities will remain maximally indeterminate.

⁶¹ See, e.g., Isaac Levi, The Enterprise of Knowledge 296-98, 399-423 (1980). See also Isaac Levi, "Induction as Self Correcting According to Peirce," in Science, Belief and Behavior 127 (Hugh Mellor ed., 1980).

⁶² Bayesian Conditionalization: Let $p(\cdot)$ be a X 's probability function and $p_e(\cdot)$ be the probability function which results when X learns that e is true and incorporates it into his background theory. Then $p_e(h) = [p(e|h) * p(h)] / p(e)$.

To prove Levi's assertion, assume for the purposes of contradiction that Tony's probabilities are maximally indeterminate before he learns his doctor's recommendation, call this report e , but not after. That must mean that Tony rules out some probability function $p_e(\cdot)$. Let $p_e(h) = x$. However, $p_e(\cdot)$ can be obtained via Bayesian Conditionalization from the function $p(\cdot)$, where $p(h) = x * p(e)/p(e|h)$. We know that Tony does not initially rule out this function because his initial credal state is maximally indeterminate. Hence, when Tony starts with maximally indeterminate probabilities, he will end up with maximally indeterminate probabilities if he updates his credal state via Bayesian Conditionalization.

All is not lost. Tony can decide ahead of time to constrain his future self to assign a probability .95 to h when the report e is made. As Levi describes it, Tony does not treat e as "evidence" which he incorporates into his background theory and off of which he conditionalizes. Rather, he treats it as a causal "input," and constrains himself to follow through on instructions to which he has previously committed. Because of his doctor's reliability, this "up front" choice not to accept e as evidence later maximizes epistemic utility.

Levi's argument, therefore, demonstrates that the Decision Model cannot account for the normativity of authoritative advice. For it would be irrational for someone in Tony's situation to accept the recommendation of an expert as true absent a constraint to do so. In "sequential" play, agents are required to update via Bayesian Conditionalization and, therefore, those with indeterminate probabilities cannot harness the information value of authoritative recommendations.

Yet, Levi's argument can be used to support Raz's claim that authoritative advice differs in kind from ordinary advice. In situations where ordinary advice is normally solicited and given, we are not so ignorant that our probability judgments are maximally indeterminate, or nearly so. We, therefore, are at

liberty to treat ordinary advice as evidence and update via Bayesian Conditionalization. However, treating authoritative advice as evidence destroys its value, because in these situations our credal states may be highly indeterminate. It is only by constraining our future selves that we can learn from experts.

B. Compensating for Shortfalls in Rationality

In some situations, authorities are instrumentally valuable because they save us from having to engage in costly and risky deliberation. We can simply follow their pronouncements and be reasonably confident that we will be making the right selection most of the time. What happens when authorities give the wrong answer? According to the Decision Model, it may still be rational for the agent to decide to follow their recommendation. Directives that cut down on deliberation costs are likened to rational gambles -- whenever the marginal expected costs of deliberation exceed the marginal expected benefits, an agent should rationally choose to follow directives in every applicable case. Losing a bet does not indicate irrationality if it was rational for the gambler to have taken his chances in the first place.

Consider the following example. Liz hates deciding what to order when in a restaurant. She also thinks that waiters' culinary judgments tend to be as reliable as her own. She figures, therefore, that it would be better for her, all things considered, to treat their evaluations as authoritative rather than to agonize over what to eat.

According to the Decision Model, each time Liz follows through on a waiter's recommendation, she is making a rational decision. Even though Liz will pass up a terrific dish every so often by following the waiter (let's say that she believes that the waiter will mention a great special 5% of the time), it is

nonetheless an acceptable risk given that the probability of such an occurrence is low relative to the savings that she can expect to accrue by not deliberating.

The problem with this reasoning is that it's fallacious: once Liz knows the specials, then the probability that terrific dish is being offered is no longer .05 -- it is either 1 or 0. Probabilities are always computed relative to the total evidence available to the agent and, relative to the evidence available to Liz, either there is a terrific dish being offered or there isn't. Either deliberation is worth it or it isn't. She is no longer engaged in decision making under risk -- she is facing a decision problem under certainty. The waiter's recommendation is useless to her now, because the only way she can know whether to follow the directive, i.e., whether it gives the right answer, is to deliberate. If the waiter gives the wrong advice, then it is irrational for her to decide to apply it.

Lest there be confusion on this matter, I am not claiming that the total evidence requirement requires Liz to deliberate. It merely requires that her choice be one that would be endorsed by ideal deliberation on all the available evidence. Liz is now in a pickle: the only way she can know whether to deliberate is to deliberate. At this point, the principles of rationality offer no guidance; whatever she decides to do is not irrational.

If, on the other hand, Liz were constrained to follow the waiter's recommendation, the probability that it pays to deliberate is irrelevant. Liz will be forced to conform regardless of the probabilities of success *ex post*. This constrained behavior is optimal from an *ex ante* perspective and that is why we should see her behavior as rational.

Once again, the Decision Model is unable to account for the normativity of authoritative directives. When agents have open minds, they are required to maximize expected utility. Only by committing

oneself in advance to accept the direction of an authority can that direction be a reason for action.

C. Coordination

According to the Decision Model, authoritative directives affect practical reasoning in a coordination game by providing evidence about the activities of the other players. Authoritative directives are capable of providing evidence of others' activities because they create salience. The rule requiring me to stop at a stop sign, for example, makes the joint strategy "my stopping, crosswise traffic not stopping" salient. Hence, the rule constitutes good evidence that crosswise traffic may be crossing the intersection, making it rational for me to decide to stop.

I would like to challenge this analysis. First, I will argue that authorities do not solve coordination problems by creating salience. Salience is one way of solving a coordination problem, authority another. Second, I will try to show that, assuming that players commonly believe that each other are rational, authoritative directives could not constitute evidence of others' activities and hence could not solve coordination problems. If authorities are able to solve coordination problems, the Decision Model cannot be correct.

To see why salience is the wrong concept to employ with respect to authoritative solutions to coordination problems, it would be best to consider briefly the reasons that the concept of salience was first introduced. In the Strategy of Conflict, Thomas Schelling was interested in correcting the dominant tendency in game theory that viewed all instances of conflict as cases of pure conflict. As he pointed out, many important strategic situations involve non-zero-sum components. It is often in the parties' interest to cooperate. The rub is that in many strategic situations, explicit bargaining is not feasible. Because the

existence of conflict frequently leads to a breakdown in communication, the mutual selection of a coordination equilibrium cannot be had by simple verbal manifestation of intention. A meeting of the minds must be had by a process of imaginative second-guessing, each trying to figure out what the other agent would expect him to select knowing that other knows that he knows this.⁶³

The main work to be done in tacit bargaining, therefore, is guesswork. Each side tries to guess what the other will guess. The option which possesses the property of "being most easily guessed" Schelling dubbed the "salient" option. Salience is, therefore, a function not so much of the uniqueness of an option, but rather its conspicuousness. Each side must know that the option stands out in some respect, drawing and focusing attention, and each party must know that the other knows this.

To call tacit bargaining an exercise in guesswork is not to claim that the parties cannot be fairly confident that coordination will be achieved. Experimental studies have shown that humans are remarkably good tacit bargainers. The point is that coordination of expectations in these cases is not achieved by a process of communication. Each side must "guess" at what the other is thinking, given that there has been no explicit revelation of intention.

It is crucial to note that salience calculations are required as a compensation for the lack of communication. "The concept of 'coordination' that has been developed here for tacit bargaining does not seem directly applicable to explicitly bargaining. There is no apparent need for intuitive rapport when

⁶³ One of Schelling's examples of a "mixed-motive" game involves the case of chemical warfare. Given that the nations fighting in World War II all had an interest in limiting the number of casualties that would result, some mutual restraint on the use of nerve gas was to known by all parties to be desirable to all parties. Yet, the number of possible prohibitions were considerable, e.g., "No Gas", "Some Gas", "Gas only used on military personelle", "Gas only to be used in self-defense", etc. Communication between the various warring factions was impossible at the beginning of the conflict and testing of the various options could not be had, considering that one failure of coordination could have prevented subsequent cooperation (it's hard for an army to restrain itself from gassing an army who has just gassed it). Tacit bargaining resulted. Of all the rules that could be selected, the absolute proscription of the use of nerve gas was the simplest. Each side conjectured that the other would pick the rule that the other could guess they would guess, and since one rule was the most conspicuous in its simplicity, it served as a focal point for agreement.

speech can be used; and the adventitious clues that coordinate thoughts and influenced the outcome in the tacit case revert to the status of incidental details.⁶⁴ It is a mistake to say when two parties agree to follow a given course of action, they each act on the same option because that option is now most salient. In these situations, salience is otiose.

Coordination via authority shares some properties with tacit bargaining, others with explicit bargaining. Like tacit bargaining, there is no direct communication between the parties wanting to coordinate. The authority acts as a "one-sided" signaling mechanism, i.e., signaling is performed in order to coordinate the actions of members of the audience, whereas in explicit bargaining the communication is "two-sided", i.e., the transmission is meant to coordinate the actions between communicator and audience.⁶⁵ Yet, like explicit bargaining, there is a pre-existing procedure for knowing which option the other will exercise. In authority-based cases, agents act on the option announced by the authority; in explicit bargaining, the agents act on the options agreed to. In tacit bargaining, contrawise, the agents do not initially know which option they ought to select. They do look for the most salient option, but which option is most salient is to be determined by guesswork ("what does the other person think is the most conspicuous option given that she knows that I am trying to figure that out and vice versa?"). There is no mechanical algorithm for selecting one coordination equilibrium which is common knowledge.

Just as it is a mistake to talk about salience in the context of explicit bargaining, so to is it when discussing authoritative direction. Authority is justified conventionally because it remedies the problems of non-communication. Given the obvious transaction costs associated with explicit bargaining, many

⁶⁴ Schelling, The Strategy of Conflict, 67 (1960).

⁶⁵ See David Lewis, Convention, 122, passim, for a discussion of the differences between one- and two-side signaling mechanisms .

coordination problems have to be solved absent two-sided verbal interaction. To overcome this, the authority acts as a one-sided signaling mechanism, allowing the parties to solve their coordination problem through the use of a shared algorithm. An authority no more makes an option salient, than do two people who agree to select one course of action make that course of action salient. In both cases, the parties have a settled technique that allows them to effortlessly solve a problem that, absent such a procedure, would be rather more difficult, if not impossible, to circumvent.

The argument against the Decision Model is not simply that its supporters have used an inappropriate word, i.e., "salience", but rather that saying that authority sometimes or always creates salience masks the unique way that coordination problems can be solved in such cases. First, and most importantly, when authorities are involved, the answers to coordination problems are settled in advance. Authorities take the guesswork out of coordinating behavior. Second, considerations that would have been relevant to establish salience become irrelevant in authoritative contexts. Authoritative solutions, therefore, tend to be more stable than their salient counterparts given that directives are not invalidated by changes in their content's conspicuousness.

Supporters of the Decision Model might accept this objection and admit that authoritative directives do not create salience. Yet, they might nevertheless argue that, however authoritative directives focus attention on equilibria, the fact that they do focus attention gives each person a reason to believe that others will choose to act in accordance with the equilibria. Authoritative directives are capable of solving coordination problems because they can sometimes constitute strong evidence that others will conform to the directives and thus making it rational to choose to conform as well.

However, this still will not do. For if we assume that the players are rational and commonly

believe in each other's rationality, the Decision Model does not yield the result that it would be rational for any player to follow such directives. Consider the following chain of reasoning: The Decision Model claims that it is rational for a player, call him X, to decide to comply with such a directive when, and only when, X can establish that it provides good evidence about other players' behavior. But the directive provides good evidence about others' behavior only when it would be rational for others to follow it. However, if these players are rational, the question of whether it is rational for them to follow the directive is the same as whether it is rational for X to follow the directive. Hence, X can establish that the directive constitutes good evidence only if X can first establish that it is rational for him to follow the directive.

X has now traveled in a circle. If X want to establish the rationality of his following the directive, it seems that he must already know that it is rational for him to follow the directive. But since he is trying to establish the rationality of following the directive, he cannot assume the proposition for the purpose of proving it. So, if X does not already believe that it is rational for him to follow the directive, he will never come to that conclusion.

The conclusion I think that we should draw from this argument is that coordinating rules cannot solve coordination problems if the players are rational and it is commonly believed that every player is rational and treats the directives simply as evidence about the other players' behavior. The players must believe that at least some of the other players either (1) treat the directives as constraints on action or (2) believe that at least some of the other players treat them as constraints on action. The fact that some players are committed to the authority ensures that they will follow the rules and these commitments, or at least the beliefs that some are so committed, will enable the appropriate expectations to form so that coordination may take place.

D. Feasibility

The central, and as yet unanalyzed, concept of the Constraint Model is, of course, “feasibility” – disregarding a directive is seen as “infeasible,” as not being an “option” available to the agent. It is, therefore, important that I say at least a few words about the concept of “feasibility” that the Constraint Model employs.

We will say that a course of action is a feasible option for an agent if and only if that agent has the ability to perform that action for a reason. A course of action is infeasible, therefore, whenever the agent cannot perform the action for a reason -- that is, when that action is not a possible intentional action.

Now, if we were to speculate about the mechanisms by which authoritative directives render disobedience infeasible, two possibilities come to mind. First, submitting to an authority might be the psychological counterpart of precommitment, where the agent is physically disabled from moving his body in certain ways. This physical disability would not be imposed by some external device, but would be generated from “within” by purely psychological means. Just as we can get our hand to move by intending to move it, we can get it to stay put by intending not to move it.

I am inclined to think that this approach is unpromising, at least if we attend to the phenomenology of rule-guided behavior in general. One normally does not think of rule-guided behavior as involving sheer physical constraints. When someone adopts a rule never to take another drink of alcohol in a bar, he does not think that somehow he will be precluded from physically ordering the drink. It is not as if he thinks that if he opens his mouth, he will not be able to utter the words “I’ll have a beer” or if he does get

a beer, his hand will be frozen in the down position. In this respect, he is unlike Ulysses: Ulysses can intend to break free of the rope and try to do so. Rule-guidance, on the other hand, seems more like a constraint on the will, on intention-formation. If John does intend to drink alcohol, he will probably end up doing so.

I would like to suggest instead that authoritative directives normally prevent agents from “willing” disobedience to its demands, not by physically disabling the will’s expression. How might these constraints on the will arise? First, the directive, coupled with the commitment to the authority, might prevent the agent from considering the reasons for disobeying. The reasons for disobeying would, in other words, be “repressed” by the directive. This repression blocks the possibility of intentional action contrary to the directive -- the agent, being unaware of reasons for not complying, would be unable to disobey for a reason. Second, the directive, coupled with the commitment to the authority, might disable the person’s normal psychological inhibitions. An agent in the grip of an authority might no longer be able to withstand certain emotional pressures, such as guilt and shame, even though the agent would be aware of the important reasons for not complying.

All of this is speculation, of course, but I do think it is plausible speculation. If one thinks about the abilities of the human mind, undoubtedly one of its powers is that of repression. We normally hide “reasons for action” from ourselves all the time. Awareness of all the reasons we had for acting one way rather than another would simply be too painful to bear. I am suggesting that when we submit to authority in order to benefit from them instrumentally, we harness this ability, at least sometimes, to repress certain facts and wishes from our practical reasoning. With respect to rules, the most common expression indicative of this process is: “I have a rule against do this so I am not even going to think about acting

otherwise.” By refusing to think about it, one guarantees that one cannot break one’s rules for reasons.

Most of the time, however, we aren’t even aware of repressing reasons for violating our commitments. We simply follow rules and directives without any struggle or conflict, even in situations where if we thought about it, we would be deeply disturbed by our actions. They become almost like habits in that they can be executed virtually without thinking.

The ability to repress reasons is not one that everyone has nor do they have it to the same degree. Nor is the need to repress reasons as important in some people as in others, chiefly because all people do not monitor their actions with the same frequency. As control theorists say, some have longer “feedback loops.” In situations where the need for guidance is most acute, for example, in the military, people need to be trained to lengthen their feedback loops. Boot camp consists in training people to ignore their instincts, to react first rather than think. This is accomplished partly through the complete regimentation of life and partly through the compelled performance of absurd tasks like cleaning floors with toothbrushes. The easiest way to get people to repress reasons is, therefore, to stop them from prompting themselves for reasons in the first place.

As for disinhibition, it is well-known that people who act subject to orders can be made to act in monstrous ways. It is not that they don’t realize what they are doing; rather, they seem to be unable to oppose the demands of the authority issuing the directives. In reporting on the results of his famous obedience experiments in the 1960’s, Stanley Milgram vividly described this phenomenon. (Recall that in these experiments, Milgram wanted to see how far ordinary people who go in hurting others when being ordered by an authority to do so. The subjects were told that they were to participate in an experiment on learning and negative reinforcement. Whenever the supposed learner would make a mistake, the

subject was directed by the leader of the experiment to administer an increasing series of electric shocks to punish the learner for the mistake. To Milgram's surprise, many of the subjects followed the leader's orders and administered what the subjects thought were extremely painful shocks. What was even more surprising is that the subjects often expressed great reservations about continuing with the experiment but followed through anyway because they were ordered to do so.) This is Milgram's description:

Some people interpret the experimental situation as one in which the subject, in a highly rational manner, can weigh the conflicting values in the situation, process the factors according to some mental calculus, and base his actions on the outcome of this equation. Thus, the subject's predicament is reduced to a problem of rational decision making. This analysis ignores a crucial aspect of behavior illuminated by the experiments. Though many subjects make the intellectual decision that they should not give any more shocks to the learner, they are frequently unable to transform this conviction into actions. Viewing these subjects in the laboratory, one can sense their intense inner struggle to extricate themselves from the authority, while ill-defined but powerful bonds hold them at the shock generator. One subject tells the experimenter: "He can't stand it. I'm not going to kill that man in there. You hear him hollering in there. He's hollering. He can't stand it." Although at the verbal level the subject has resolved not to go on, he continues to act in accord with the experimenter's commands. Many subjects make tentative movements toward disobedience but then seem restrained, as if by a bond.⁶⁶

Milgram hypothesized that by accepting the authority of another, one relinquishes the ability to inhibit actions in light of one's own values: "The inhibitory mechanisms which are vital when the individual element functions by itself become secondary to the need to cede control to the coordinating component."⁶⁷ He subscribed, in other words, to the Constraint Model.

In some sense, it is irrelevant whether the mechanisms I have described are present in normal functioning humans. Strictly speaking, authoritative directives can serve the benefits they are meant to

⁶⁶ Stanley Milgram, *supra* note 60, 148-49.

serve just in case we think that we have the ability to constrain our future selves. Whether we can actually do such a thing is another matter. We might be wrong in our belief that authorities can give rational agents reasons for action, but this belief is neither incoherent nor implausible.

E. Interim Conclusion

In the previous sections, I argued that the Decision Model cannot account for the instrumental potential that authoritative directives hold for rational agents. I also argued that the Constraint Model does provide an adequate explanation. Rational agents can benefit from the directives provided them only if they do not make choices about whether to obey individual directives. They may make choices to commit themselves to the authority – but those choices are choices not to make future choices.

Moreover, I suggested that the Constraint Model can provide an adequate resolution to the paradoxes of authority. When authoritative directives give the wrong results, it is neither irrational nor immoral for the committed subject to comply, given that compliance is the only feasible option. When authoritative directives give the right results, the directive is relevant to practical reasoning insofar as it affects the feasibility of non-conformity.

However, this argument is not sufficient to end the discussion of the paradoxes. For even if one were to accept the truth of the Constraint Model, the philosophical anarchist might still argue that the paradoxes remain. The Constraint Model solves the paradoxes only when a particular theory of legitimate authority is presupposed, namely, one that embraces the Normal Justification Thesis. If the

⁶⁷ *Id.* at 129.

legitimacy of an authority were not dependent on its ability to generate instrumentally valuable directives, the paradoxes would remain unsolved. For the Constraint Model has nothing to say about obedience that is not motivated by the desire to harness the instrumental potential of directives. It is consistent with the Constraint Model that individuals who obey directives despite their lack of instrumental value are making decisions to obey. If so, the paradoxes remains: how can it be reasonable/morally acceptable for agents to choose to obey directives that are mistaken?

The paradoxes remain, of course, only if it is possible for an authority to be legitimate despite its inability to generate instrumentally valuable directives. As I will argue in the next section, we must take this possibility seriously.

VII. Serving the Governed

Despite all of the critical attention focussed on Raz's theory of authority, few have challenged the validity of the Normal Justification Thesis. Raz's critics, by and large, have accepted his claim that authorities are legitimate just in case subjects are likely to do better in terms of the reasons that independently apply to them by deferring to their directives than if subjects attempted to conform to these reasons directly. The main bone of contention, as we have seen, has centered on the Preemption, not the Normal Justification, Thesis.

The appeal of the Normal Justification Thesis, I believe, stems from its being an expression of the Service Conception. According to the Service Conception, it is the function of authorities to serve the governed. The Normal Justification Thesis regards this service as consisting in the provision of directives

to subjects so that they might better conform to the reasons that bind them.

The idea that authorities are in our service, not vice versa, is one of the most important ideas to have emerged from the Enlightenment.⁶⁸ Enlightenment thought rejected the notion that authorities derive their power from superior birth or social status. Indeed, it has been the burden of modern political theory to explain how authorities can have the power they claim to have despite the fact that no person is “better” than anyone else. The Service Conception provides the standard response to this dilemma: authorities have the power to tell us what to do because we benefit, in some sense, from their having such power.

As I will argue, however, the Normal Justification Thesis is but one expression of the Service Conception. Accordingly, I will distinguish between two different models of service, one that sees the function of authority as mediating between reasons and persons, and another that understands them as arbitrating between rival parties. I will argue that the traditional liberal understanding of service consists not in mediation but in arbitration. Rather than capturing a dominant theme of traditional liberal thought, the Normal Justification Thesis represents a somewhat radical understanding of the function that legitimate authorities are meant to serve.

It is not, of course, an argument against the Normal Justification Thesis that it breaks with tradition. Rather, I will argue that Raz’s theory of authority is flawed because of the inadequate justificatory role that it accords to democratic decision-making. According to the Normal Justification Thesis, the value of various schemes of power-sharing in a society is understood primarily in instrumental terms – one structure of government is more legitimate than another when one is more likely to track the

balance of dependent reasons than another. Consequently, democratic structures are preferable to undemocratic ones if and only if the former generate “better” directives than the latter.

However, I will suggest that the value of democratic decision-making does not lie in its instrumental value. Rather, democratic procedures legitimate authority because they represent power-sharing arrangements that are fair. In democracies, every citizen is given an equal ability to exert control over their life and the life of the community. Rather than violating one’s autonomy, heeding rules that one believes to be mistaken can be an affirmation of the value of autonomy in general. It shows respect for the rational faculties of others, recognizes the fairness of accepting burdens in cooperative ventures and supports the equality in distribution of power through society.

A. Mediation and Arbitration

Authorities can serve their subjects in one of two ways. First, they might serve their subjects through the guidance that their directives provide, i.e., by enabling subjects to achieve benefits that they would not have been able to achieve without the directives. In this essay, we have seen the many benefits that authoritative guidance can secure via its directives, i.e., dispense valuable information, compensate for cognitive shortcomings, economize on deliberation costs, combat weakness of the will and coordinate behavior.

Second, authorities might serve their subjects by providing them with a way to resolve their disputes on normative matters. Disagreements between parties can be settled by appealing to the

⁶⁸ Or, as I should say, “reemerged,” as the Service Conception was an important theme in classical political thought.

authoritative determinations of a certain person or persons, such as clergy, teachers, parents, officials, courts, legislatures, or agencies. On this account, success is not measured by the content of the directives issued. Rather, authorities serve their function when the directives issued are capable of resolving actual or potential disputes.

Accordingly, we can distinguish between two service conceptions of authority. The first, which might be called the ‘Mediation Model,’ understands the function of authority to be the mediation between reasons and persons. Authorities are legitimate for a subject to the extent that authorities serve this function effectively, i.e., the subject is better off in terms of the reasons that bind her by complying with directives provided than if she attempted to conform to those reasons directly. The chief proponent of the Mediation Model is, of course, Joseph Raz.

According to what might be called the “Arbitration Model,” the function of authority is to act as an arbitrator between subjects. Authorities are legitimate for a given subject just in case the acceptance of the process as binding by some of the parties generates a moral obligation for the subject to abide by the outcome. The type of acceptance, the parties that must accept the process and the nature of the moral obligation generated by such acceptance will vary depending on the type of Arbitration Model. A social contract theorist, for example, would understand the acceptance as an act of consent by the subject and the obligation generated to be a promissory one. A fair play theorist, on the other hand, would understand the acceptance as the willing receipt of the benefits of the process by the subject and the obligation generated would be one of fairness, i.e., that parties shoulder the burdens of a process when they also willingly accept the benefits from it as well.

The Mediation and Arbitration Models differ in three respects. Most obviously, they differ in the

main function attributed to authorities. For the Mediation Model, the function is one of mediation between reasons and persons; for the Arbitration Model it is of arbitration between rival parties. Second, these models differ in the relationship they draw between the function of authority and its legitimation. In a Mediation Model, the relationship between function and legitimation is direct: authorities are legitimate for a subject if and only if they serve their mediating function for that subject. In an Arbitration Model, by contrast, authorities are not necessarily legitimate for a subject simply because they arbitrate disputes involving that subject successfully. The connection is more indirect: the arbitration function gives parties reason to accept the outcome of the process and it is this acceptance, not the successful performance of the dispute-resolution, that lends legitimacy to the process.

Third, the models differ in terms of the ultimate grounds of legitimation. In the Mediation Model, authorities are legitimated by dependent reasons. When facing legitimate authority, each subject is likely to do better in terms of the reasons that bind him if he complies with the directives than if they did not comply. In the Arbitration Model, subjects might be less likely to do better in terms of the dependent reasons. What binds subjects is the acceptance of the process as binding by some of the parties.

The contrast between the two models might be summed up as follows. In the Mediation Model, obedience itself is instrumentally valuable. In the Arbitration Model, the parties do not benefit through their obedience. Obedience, rather, is the moral price that parties must pay in order to secure the compliance of others.⁶⁹

⁶⁹ In addition to these two accounts, we might imagine a third mixed model of authority. The “Mixed Model” is a disjunctive combination of the Mediation and Arbitration Models. It legitimates authority just in case it either successfully mediates between reasons and persons or the commitment on behalf of the subjects to abide by the

B. Mediation and Democracy

The Arbitration Model has surely been the dominant account of authority in modern liberal theory.

The classical liberal theorists such as Hobbes, Locke and Kant, all believed that the foundation of legitimate authority lay with their ability to arbitrate disputes. The move from the state of nature to civil society, they argued, was necessitated by the costs associated with anarchy, i.e., with the absence of a person or persons to which feuding parties could appeal in order to resolve their disputes. It is the function of authorities to serve their subjects, but not primarily by issuing instrumentally valuable directives, but rather by issuing directives in the first place. Ideally, of course, the directives issued should be morally appropriate and conducive to the common good. Yet, the obligation to obey these directives does not depend on their meeting, or even coming close, to this ideal.

Despite its fringe status in modern liberal theory, the Mediation Model has found a toehold and is gaining strength. Unfortunately, this is not the occasion for a full investigation of the Mediation Model as compared to the Arbitration Model. I would, however, like to suggest that the Mediation Model might be less plausible than has hitherto been thought.

According to the Mediation Model, the legitimacy of authority is determined exclusively by its ability to provide instrumentally valuable directives. The origins of the authoritative regime seem to be irrelevant, at least from the standpoint of the obligation to obey. Institutions, as we have seen, are measured primarily in instrumental terms – one structure is more legitimate than another just in case the former leads to more effective mediation than the latter.

determinations of authority actually generate a moral obligation to obey.

However, this instrumental conception ignores the intrinsic value of democracy. The legitimacy of rule is generally not judged exclusively, or maybe even primarily by its output, but rather by its input, i.e., by whether the regime has been determined, and is supported, by the populace. Platonic political theory has been derided for centuries precisely because of its top-down structure. Despite the philosopher-king's excellent mediation skills, his right to rule is defective unless his power has been affirmed by those over whom it is exercised.

The Arbitration Model, by contrast, has room to accommodate the importance that democratic rule plays in legitimating authority. To see this, we must return to the paradoxes of authority.

C. Arbitration and the Paradoxes of Authority

The Mediation Model attempts to solve the paradoxes of authority, as we have seen, by taking a thoroughly instrumental approach to authoritative directives. This strategy will not work for the Arbitration Model, however, as the normativity of directives does not depend on their instrumental value. The Arbitration Model sanctions obedience, for example, even when the directive is not supported by the balance of content-dependent reasons, subjects are aware of this and have the ability not to comply. We must see, then, whether obedience to authority can be rendered compatible with both rationality and autonomy.

Let us begin with the paradox of authority and rationality. Assume that a legitimate authority issues a directive that is supported by the balance of content-dependent reasons. Does the directive give a subject a reason for action? Yes, according to the Arbitration Model. Because the subject is morally

obligated to comply with the outcome of legitimate processes, he will have an additional reason to comply.

Should the subject disobey the authority, he will have committed two offenses: he will have performed an action that, independent of the directive, he should not have performed and will have violated his moral obligation to obey.

Assume now that the directive is not supported by the balance of content-dependent reasons. Can obedience nevertheless be reasonable? On the Arbitration Model, it can. Since each subject is morally obligated to comply, the directive constitutes a content-independent reason for action. The balance of all reasons, content-dependent as well as independent, might then tip in the direction of obedience, even though it would have tipped in the opposite direction had the directive not been issued.

Notice that the Arbitration Model supplies a solution to the paradox of authority and rationality regardless of whether it is true. For as long as an agent thinks it is true, it will be rational for him to obey.

By contrast, the Arbitration Model must be true in order if it is to provide a solution to the paradox of authority and autonomy. For it will not be enough merely to assert the coherence of subjects believing that they sometimes have moral obligations to comply with the results of arbitration. We would have to demonstrate that such obligations can truly exist.

Consider, for example, the social contract variant of the Arbitration Model. One might be tempted to answer the paradox of authority and autonomy by arguing that the promissory obligation engendered by consent gives agents reasons to act contrary to the balance of content-dependent reasons. However, as we saw at the beginning of the essay, this response is question-begging. If some person's will cannot give me a reason to do what ordinarily would be wrong, how can my own will give me such license? Since consent also purports to be a content-independent and peremptory reason for action, the

social contract account merely pushes the paradox one step back, namely, how can the mere fact that I have consented to abide by the results of some process give me reason to abide by the results of that process.

In what follows, I will suggest that the moral obligation to obey authority can be generated under certain conditions in a democratic polity. Roughly, the idea is deference to democratically elected authority is deference to a power-sharing arrangement that is socially necessary, empowering and fair. By disobeying, subjects are unilaterally, and hence unreasonably, setting the terms and direction of social cooperation.

A sketch of this argument begins with the truism that social cooperation is not, as a practical matter, possible without the availability of procedures for the resolution of conflict. Disagreements between parties as to the appropriate terms of social interaction and the division of social surplus would either forestall or derail individual and joint pursuits. And absent acceptable resolution, disputes would fester into outright feuds and internecine battles would likely threaten the very survival of the community.

The next step in the argument involves the claim that democratic institutions are empowering institutions. In democracies, citizens are granted the power to exert control over their lives by allowing them, through the franchise, to affect the terms of social cooperation and the direction of collective pursuits. They may affect the shape of the social landscape either directly, by plebiscite, or, more familiarly, indirectly, by the election of representatives. The protection of free speech also enables citizens to influence the social structure and objectives by permitting them the opportunity to persuade their rivals, and the uncommitted, of their views. As opposed to oracles, trials by ordeal, or coin flips, therefore, democratic procedures allow individuals some input into the resolution of their disputes. Democracies

give expression to, and create opportunities for the exercise of, the individual's autonomous capacities.

Lastly, democratic processes constitute fair procedures for the resolution of disputes. Democratic processes are fair because power is shared in a roughly equally manner.⁷⁰ Equal power-sharing consists, first, in the equal voting power that individuals, or groups to which they belong, have in the selection of policies or election of representatives. Second, equality of power is determined by the equal opportunity that individuals, or groups to which they belong, possess to express their views and to persuade others as to the value of their positions.

The crucial step is the claim that it is unreasonable for an individual not to abide by socially necessary, empowering and fair procedures. To motivate this assertion, consider the objection that might offer to challenge it.⁷¹ It might argue that it is an unreasonable infringement on personal liberty to require someone to be bound by a procedure they did not voluntarily accept. No one has the right to demand that someone shoulder burdens when those benefits have been thrust upon them.

This objection, however, lacks merit. It ignores the fact that personal liberty has value only when a scheme of social cooperation is already in place. One cannot complain that one's ability to pursue projects in the manner one sees fit would be overburdened when the ability to pursue those projects essentially depends on everyone else's restraint. The very assertion of personal liberty indicates that the objector willingly accepts the benefits of such procedures. The objection, in other words, wants its cake and to eat it too.

⁷⁰ "Democracy" is being used here in a moral, not sociological, sense. A procedure that does not involve equal power-sharing is, therefore, not a democratic procedure, even if everyone has an equal right to vote. The United States might not be a true democracy in this sense, because individuals, or the groups to which they belong, do not have equal voice in public debate.

⁷¹ See, e.g., Robert Nozick, Anarchy, State, and Utopia, 90-96 (1974).

One who disregards the outcomes of a socially necessary, empowering and fair procedure, we might say, acts like a dictator: both unilaterally “dictate” the terms of social interaction to others and thereby exercise inappropriate control over the lives of his fellow citizens. It is no defense for the rebel to point out that the procedure produced an incorrect result – for whether it did or not, it is not “up to him” to impose his own judgment on others.

The conclusion I would like to draw is that disobedience to the democratic will, at least under certain circumstances, amounts to an unreasonable arrogation of power. Those who act in such unreasonable manners deprive those in the majority of three important goods. First, they deny the majority the outcome that they had a right to expect. Second, they deny the majority the control over their lives that they had the right to exercise. Third, they deny the majority the respect due to them as equal participants in a fair power-sharing arrangement.

Somewhat surprisingly, it turns out that submission to authority, rather than leading inexorably to the violation of autonomy, actually manifests respect for autonomy, understood here as the power to control one’s life. Deferring to democratically elected authority or selected policies is deferring to one’s fellow citizen. In doing so, one pays respect to the importance that people are allotted a certain control over their lives and the fairness of sharing that power equally.⁷²

⁷² That autonomy plays a significant role in grounding the obligation to obey democratic procedures suggests that their authority is also limited. The obligation would lapse whenever democracies failed to respect the autonomy of their citizens. This would occur, for example, when democratic control is exercised over purely self-regarding activities. The obligation to obey the outcomes of democratic procedures would also expire when democracies systematically disempower certain groups through such outcomes.

VIII. Conclusion

Let us return to the dispute between Eliezer and the Rabbis. Was Eliezer right not to submit to the mistaken judgment of the Rabbis? I think it is fair to say that Eliezer acted badly. If he truly cared about the Will of God, he would have submitted, for God had willed that legal decisions are to be decided by majority vote.

Eliezer manifested a vice that is not uncommon among the pious, a vice which might be called “excessive purism.” Excessive purists always insist on acting in the technically right manner. They refuse to corrupt themselves, to dirty their hands by descending to the level of the lumpenproletariat and to act as the benighted do. However, as I have tried to argue, one can have reasons to abide by the will of another, even when one knows that they are wrong. To loftily stay above the fray can manifested extreme disrespect for one’s fellow citizens.

Perhaps the dangers of excessive purism in matters of authority are slight compared with the vices of the opposite extreme, namely, moral apathy, sloth and servility. A sceptical attitude towards authority is perhaps the healthiest stance to take. But such scepticism, I have tried to suggest, can go too far.