Title
Legitimacy’s Stutter: The Philosophical Anarchist Challenge and Deliberation’s Uncertain Solution

Permalink
https://escholarship.org/uc/item/2sm1z2q2

Author
Ward, Alan

Publication Date
2017

Peer reviewed|Thesis/dissertation
Legitimacy’s Stutter: The Philosophical Anarchist Challenge and Deliberation’s Uncertain Solution

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Political Science

by

Alan Ward

Committee in charge:

Professor Tracy B. Strong, Chair
Professor Gerry Mackie, Co-Chair
Professor Harvey Goldman
Professor Marcel Henaff
Professor Andy Lamey

2017
The Dissertation of Alan Ward is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

_____________________________________________________________________

Co-Chair

Chair

University of California, San Diego

2017
DEDICATION

Writing a dissertation is hard. I don’t know anyone who successfully wrote one without sometimes feeling as though they were running into a brick wall. Over. And over. And over. But some people made the process less difficult. These are the people who put padding on the wall, bandaged my forehead, and encouraged me to keep banging. This dissertation is dedicated to them.

First, thanks go to my committee – Tracy Strong, Gerry Mackie, Harvey Goldman, Marcel Hénaff, and Andy Lamey. The breadth and depth of their knowledge baffles me. At this point I just take for granted that when they speak, whether in a group setting or one-on-one, I’m about to hear something perceptive and liable to force me to rethink some of my positions. Thank you for your professionalism, your help, and for making me smarter. The same goes to Prof. Isaac Balbus, who supervised my Masters thesis at the University of Illinois at Chicago. Ike taught me that it’s possible to be critical while being kind.

I also want to thank my grad student cohort and the friends I made in San Diego. My friends Kathryn and Rhys fed me, laughed with me, and made sure I never had to have a pint alone. Rick Malabo helped me clear my jumbled thoughts by going on frequent walks and showing me unrelenting kindness. Marc Rolnik challenged me, but did so as a true friend, providing wine and chocolate as we debated politics. Konstantin Ash entertained and one-upped my ridiculous assertions until our talks reached the level of sublime absurdity. Thanks to Simone Günther, who taught me many lessons, including humility by tying me in a pizza-eating contest. Finally, my relationship with Rachel MacCratic has been a major influence on my thought. She is one of the most outsized and original personalities I met in any field, and I can’t imagine my time in San Diego without her.

Thanks go to Dr. Andrew Sward for all the laughs and rants. Sara Wong supported me during the bad times and celebrated with me during the good ones. I cannot thank her enough for being a part of my life. Peninah Rosenblum: you have filled my life with joy and have encouraged me to be a better man. I’m so happy you’re my partner.

Thanks to Mark Bigney, whose brilliance and almost frightening uprightness has made him my stock character when discussing intelligence or Kant’s categorical imperative. Luiza Axioti has been a model of integrity and courage. Marc-A has shown me what it means to combine compassion, an analytic mind, and an assertive drive for novelty. In a very real sense, I never would have gone to graduate school had it not been for Marc Peridis, Pirada Chamandy, and Jennifer Teoh. Thank you.

Erat, Darc, Jeron: seeing you is always a highlight when I visit Montreal. It wouldn’t be home without you. Thanks for the hospitality and, most of all, the ridiculousness.

Finally, this work is especially dedicated to my family. Andrea and Jason: I love and admire both of you. Simone and Gwen: I couldn’t ask for more wonderful nieces. You make being an uncle fun. Mom and dad, thank you so, so much for loving me and teaching me so much. I’m proud to have you as parents. <3
II.1 Affinities between consent and promising ......................................................... 110
II.2 Differences between consent and promising ..................................................... 114
II.3 Consent, promising, and informality ................................................................. 118
III. Voting ................................................................................................................. 122
III.1 Voting does not entail consent ....................................................................... 123
III.2 Social understandings of the link between consent and voting .................... 129
III.3 The non-consenting voter .............................................................................. 133
IV. Conclusion ......................................................................................................... 137

Chapter 5: Natural Duty, Dignity, and Respect .......................................................... 140
I. Existing Natural Duty Theories and response ...................................................... 141
I.1 Rawls's natural duty theory .............................................................................. 141
I.2 Philosophical anarchist response ..................................................................... 144
II. Dignity and Respect ............................................................................................. 151
II.1 Dignity .............................................................................................................. 153
II.2 Respect ............................................................................................................ 161
III. Natural duties and deliberation ......................................................................... 179

Chapter 6: Deliberative Democracy ......................................................................... 184
I. Why Deliberative Democracy? ............................................................................ 184
I.1 Problems with democracy ................................................................................. 186
I.3 Deliberation Defined ......................................................................................... 200
II. Deliberative Models: Orthodox vs. Subaltern ..................................................... 211
II.1 Orthodox Models .............................................................................................. 211
II.2 Critiques of Orthodox Models and Rise of Subaltern Deliberative Democracy .. 216
III. Critiques of Subaltern Models and Impasse ....................................................... 222
IV. Privacy and Modified Model ............................................................................. 229
IV.1 Privacy and the Internal World ...................................................................... 229
IV.2 Recap and Philosophical Anarchist-friendly Model ...................................... 234

Chapter 7: Etiquette .................................................................................................. 245
I. Politeness, Face, and Politic Behavior ................................................................. 246
I.1 Brown & Levinson .............................................................................................. 246
I.2 Watts and the Critique of B&L ......................................................................... 253
I.3 Comparing the frameworks ............................................................................. 260
I.4 Fiske .................................................................................................................. 266
II. Laughter ............................................................................................................. 274
III. Conclusions ....................................................................................................... 283

Chapter 8: Apology .................................................................................................. 288
I. The Structure of Apology ...................................................................................... 289
I.1. Background ....................................................................................................... 289
I.2. Tavuchis's Tripartite Apologies ..................................................................... 293
I.3. Apologies in Political Science ...................................................................... 305
II. Lazare and Critique of Ideal Model ..................................................................... 309
II.1. Lazare ............................................................................................................. 309
II.2. Critique of Tavuchis ..................................................................................... 316
III. Restorative Justice and Apology .......................................................................... 322
III.1. The Importance of Apologies in Restorative Justice .................................. 322
III.2. Case Study: Apologies and Braithwaite ..................................................... 334
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. Apology in Deliberative Contexts</td>
<td></td>
<td>337</td>
</tr>
<tr>
<td>IV.1 Apology and Coercion</td>
<td></td>
<td>340</td>
</tr>
<tr>
<td>IV.2. The burden of acceptance</td>
<td></td>
<td>346</td>
</tr>
<tr>
<td>V. Conclusion</td>
<td></td>
<td>353</td>
</tr>
<tr>
<td>Chapter 9: Rehabilitating Wolff</td>
<td></td>
<td>355</td>
</tr>
<tr>
<td>I. Introduction to the Conclusion</td>
<td></td>
<td>355</td>
</tr>
<tr>
<td>II. The Alienation Critique</td>
<td></td>
<td>359</td>
</tr>
<tr>
<td>II.1 Summary and description of Wolff</td>
<td></td>
<td>359</td>
</tr>
<tr>
<td>II.2 Baugh's critique</td>
<td></td>
<td>363</td>
</tr>
<tr>
<td>II.3 Response to Baugh</td>
<td></td>
<td>366</td>
</tr>
<tr>
<td>III. The Promises are Promises Critique</td>
<td></td>
<td>371</td>
</tr>
<tr>
<td>III.1 Are all promises alike? Or, how to fail at promising</td>
<td></td>
<td>371</td>
</tr>
<tr>
<td>III.2 Attenuated conception political promises</td>
<td></td>
<td>376</td>
</tr>
<tr>
<td>III.3 Promises and the social meaning of agreements</td>
<td></td>
<td>377</td>
</tr>
<tr>
<td>IV. Conclusion</td>
<td></td>
<td>381</td>
</tr>
<tr>
<td>Bibliography</td>
<td></td>
<td>388</td>
</tr>
</tbody>
</table>
VITA

<table>
<thead>
<tr>
<th>Year</th>
<th>Position/Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>Bachelor of Arts, McGill University</td>
</tr>
<tr>
<td>2007-2009</td>
<td>Research Assistant, University of Illinois at Chicago</td>
</tr>
<tr>
<td>2009</td>
<td>Master of Arts, University of Illinois at Chicago</td>
</tr>
<tr>
<td>2013-2017</td>
<td>Instructor, University of California, San Diego</td>
</tr>
<tr>
<td>2017</td>
<td>Doctor of Philosophy, University of California, San Diego</td>
</tr>
</tbody>
</table>
ABSTRACT OF THE DISSERTATION

Legitimacy’s Stutter: The Philosophical Anarchist Challenge and Deliberation’s Uncertain Solution

by

Alan Ward

Doctor of Philosophy in Political Science

University of California, San Diego, 2017

Professor Tracy B. Strong, Chair
Professor Gerry Mackie, Co-Chair

There is persistent disagreement over whether or not there is a prima facie
moral duty to obey the law. While most contemporary democratic theorists believe that such a duty exists, “philosophical anarchists” argue that such a position would constitute a violation of individual autonomy. In this dissertation, I present and assess a deliberative model of political legitimacy, one that I believe takes seriously the concerns of philosophical anarchists.

This work is divided into two parts. First, I review and update existing arguments in favor of law abidance, finding them lacking. Second, I articulate and examine the hypothesis that we (1) have a natural duty to deliberate, which in turn (2) generates a duty to abide by the results of the deliberative procedures. Drawing on literature pertaining to dignity, deliberation, politeness, and apology, I argue that deliberation is an unpersuasive answer to philosophical anarchists. The deliberative procedures can generate a duty of law abidance only if participants feel as though the process did not violate duties of respect for self and others. In order for this to obtain, the deliberators must perform actions which signal or constitute respect. Yet these concrete actions, while necessary to demonstrate respect, can serve to undermine it as well. It is the prerogative of each earnest deliberator to determine for him or herself whether or not the deliberation is respectful. If it is not, the deliberator is discharged from the political obligation to abide by the results of the (disrespectful) deliberation, although she may nevertheless have moral reasons to obey the law. What’s more, others may have good moral reasons to coerce this person into law-abidance.
Chapter 1: The Philosophical Anarchist Challenge

This dissertation is motivated by an attempt to answer the question, "Why should I obey my country's laws?"\(^1\) This motivating question, simple as it may be, circumscribes my inquiry in at least three ways.

First, the question is not why I obey the law; it's about why I should obey it. The question is moral, not empirical.\(^2\) There seems to be a rough consensus view that citizens in the US and much of Europe believe that they have an obligation to obey the law\(^3\), but this should have no bearing – at least in principle – on whether or not they have such a duty. After all, maybe well-meaning people who take this obligation for granted would change their mind if they gave the matter more thought. What's more, the question is not merely prudential. There are of course many self-serving reasons to obey the law, but I am only interested in those prudential reasons insofar as they are connected to moral considerations.

Second, the question is individuated. As this paper is an exploration of one's duties, it shouldn't matter (in principle) if there is some societal-level obligation to obey the law if this obligation does not trickle down to the individual who is trying to come to know her particular duties. The question does not ask whether, say,

---

\(^1\) To foreshadow: some of the proposed answers that I will present in subsequent chapters might explain why we ought to obey a foreign country's laws – notably when one is a visitor – but fail to explain why I should follow "my" country's laws. The status of "my" will be described shortly.

\(^2\) That said, in some cases morality may depend on empirical considerations. Most obviously, many theories will hold that we either do or might have a duty to obey if a state is at least minimally just. Less obviously, some theories – such as the estoppel (or "quasi-consent") variant of consent theory – hold that the very fact that many people believe that they have a duty to obey the law could end up changing someone else's moral duties.

\(^3\) Hurrelman et al., “Democratic-Nation State”; and Tyler, *Obey the Law*, 45-46.
Americans broadly speaking have an obligation, but whether or not a particular American reflecting on the matter has an obligation. In other words, it doesn’t matter if most Americans have an obligation to obey the law; what counts is whether or not any individual does. Perhaps most Americans do have a duty to obey the law. If that were true, I would be interested in finding out what separates those who do from those who do not.

Finally, I want to know why I should obey my country's laws. I want to make three remarks here. First, I assume that the country under discussion is at least just enough to warrant my even asking the question. There may be a discussion to be had about the obligation to obey corrupt or evil regimes, but that is not the topic of this dissertation. Second, thinking about our duties to our own country presents us with particular challenges. Normally, our duties when we go abroad are relatively straightforward. As long as the country is minimally just, we agree that going there implies agreeing to obey that country’s laws since, under normal circumstances, we could avoid the burden by deciding not to visit. We do not have this option with respect to our country of citizenship. What's more, many people believe that we have some a special relationship with our government that provides much of the thrust behind our ostensible duty to obey the law.

Third, I must clarify what I mean by “my” country and will do so in section II.2 of this chapter. For now, I want to prepare the ground by noting (or warning) that the bulk of this dissertation deals with the question of law abidance from a liberal

4 Cases such as refugees are exceptions to this.
perspective. This means that most interlocutors described in this paper accept that (1) the individual is usually the appropriate unit of analysis, (2) where coercion exists, the burden is on the coercer to justify the coercion, and (3) many laws are coercive. Taking this perspective allows me to limit the scope of my investigation while adopting a position which I believe represents a good approximation of how most Americans think about politics and political duties. It should not be taken as an actual endorsement of liberalism or an implicit repudiation of liberalism's critics.

The remainder of this chapter is organized as follows. In Section I, I introduce the debate between philosophical anarchists, who believe that there is no prima facie reason to obey the law, and archists, who believe the opposite. In Section II, I describe the two main types of modern philosophical anarchists. One the one hand we have those \textit{a priori} philosophical anarchists who, like Robert Paul Wolff, believe that the legitimate state is impossible. On the other hand, we have \textit{a posteriori} philosophical anarchists who, like A. John Simmons, believe that such a state is merely (very) improbable. Each type of philosophical anarchism requires a different kind of response, and in the bulk of this dissertation I address philosophical anarchists like Simmons.

Section III highlights tensions within Simmons' thought. Briefly: Simmons offers three criteria that an archist theory of legitimation must satisfy to persuade a posteriori philosophical anarchists that they do have a prima facie duty to obey the law. Simmons, however, does not clearly explain why some dissenters should be compelled to accept this conclusion even if the criteria are met. In Section IV I outline
some archist approaches that I do not investigate in detail in subsequent chapters but that I draw upon in order to formulate a possible response to philosophical anarchists in a later chapter. Finally, Section V concludes and presents an outline of the rest of the dissertation.

I. Legitimate Authority and Philosophical Anarchism

Another way of asking my central question is to ask whether or not a state possesses legitimate authority over me. "Legitimacy" and "authority" are broad and contested concepts, but here I will try to sidestep much of the debate and instead draw on Robert Paul Wolff's simple description in his *In Defense of Anarchism*.

Speaking of legitimate authority, Wolff says that “Authority is the right to command and, correlativevily, the right to be obeyed.” Again, this is a question of right, and all parties involved in the debate accept that it is possible that earnest citizens could fail to ascertain their actual moral rights and duties. I may mistakenly believe that I have a duty when no such duty exists or vice-versa. Someone could make the argument that a state is legitimate with respect to a citizen if that earnestly believes that it is, but this type of argument would itself require some sort of justification.

Wolff contrasts this kind of authority with the kind of expert authority wielded by people like doctors. First, he says that the doctor’s authority is only relevant to him

---

5 Wolff, *In Defense of Anarchism*, 4. Wolff frequently refers to this as "de jure authority," which he contrasts with "de facto" authority.
with regards to his health. Were this Wolff’s only distinction, he would essentially be saying that what separates expert authority and the type of authority claimed by the state is the narrowness of the former. This is not especially helpful. Even if we grant that expert authority is narrow, nothing stops someone from saying that legislators are experts at drafting laws, that judges are experts at interpreting them, and that police officers are experts at enforcing them. Taken together, these narrow fields of expertise can go a long way toward establishing a duty to obey the law, thus undermining the kind of anarchism that Wolff wants to defend. The narrowness of expert authority should not be its central distinguishing feature if Wolff’s contrast is to matter much.

Wolff’s second point is much weightier. Unlike the de jure authority claimed by the state, the authority of the expert does not entitle her to my obedience. According to Wolff, acknowledging someone’s de jure authority implies attempting to forfeit one’s responsibility for one’s actions when those actions are the result of the exercise of that authority. By contrast, Wolff wants to say that the doctor’s expertise is not of the right kind to command obedience except inasmuch as it is an appeal to reason. Unfortunately, Wolff does not develop this point at length. Thankfully, we can supplement Wolff’s account (while staying true to it) by drawing on the work of Joseph Raz.

Raz argues that the type of authority that interests Wolff should be understood as the ability to exercise "normative power," which he defines as "the ability to change

---

6 Ibid., 15.
7 Ibid., 14. Wolff does not believe that such a forfeiture is possible. Instead, he says that attempts to do so are in fact forfeitures of autonomy, which he considers fundamentally immoral.
protected reasons. These protected reasons are those which should not be affected by countervailing reasons. This becomes clearer when he later contrasts "orders" with "advice." Advice provides weight that might tip someone's decision to act toward one side or another. In some cases, the advice is so weighty that it overrides all other weighty reasons to perform or refrain from performing an action. An order, being a protected reason, is an action which overrides or, more precisely, excludes other reasons "by kind and not by weight." In other words, it makes other reasons irrelevant, not outweighed. “A request is made with the intention that it shall be taken as a reason for action and be acceded to only if it tips the balance. Orders are made with the intention that they should prevail in certain circumstances even if they do not tip the balance.”

Turning back to Wolff’s doctor, we can say that doctors usually provide medical advice to their patients, even if we sometimes refer to such advice as "doctor's orders." Speaking specifically of Wolff's example, Raz says that an anarchist "can take such instructions to be first-order reasons without conceding the legitimacy of the authority. For it is only by acknowledging that such instructions are also valid exclusionary reasons that one accepts the legitimacy of the issuing authority.” When the state claims that it is legitimate, it claims the right to give me protected reasons to perform or refrain from performing actions. The state is not a doctor.

While one need not be a philosophical anarchist to deny that the state is

---

8 Raz, The Authority of Law, 18.
9 Ibid., 22.
10 Ibid., 23.
11 Ibid.
legitimate, this dissertation is concerned with the philosophical anarchist position in particular. Although philosophical anarchism comes in many different flavors, A.J. Simmons, one of the main characters is this story, provides two necessary and sufficient conditions can be used to determine whether or not someone falls into this camp. First, philosophical anarchists believe that the state does not possess legitimate authority, that there is, as M.B.E. Smith says, no prima facie duty to obey the law. Second, philosophical anarchists believe that they have no prima facie duty to either obey or oppose the state. As Simmons puts it, they "... do not take the illegitimacy of states to entail a strong moral imperative to oppose or eliminate states" but rather that they “... take state illegitimacy simply to remove any strong moral presumption in favor of obedience to, compliance with, or support for our own or other existing states.” So philosophical anarchists believe that a state is illegitimate if it is not authorized to give commands, but they do not believe that living in an illegitimate state means that they should advocate, even in principle, the dissolution or reform of that state.

It is this second criterion that differentiates philosophical anarchists from what Jeffrey Reiman has called "political" anarchists. On Reiman’s view, political anarchists – the type of people we normally think of when we use the term "anarchist"
– typically believe that states bring about more moral "bads" than goods.\textsuperscript{16} Political anarchists seek to convince those who support the existence of states that either (1) the criteria which they – the archists – use to defend states could be better satisfied under anarchy, or (2) the criteria that archists use to defend states are mistaken.\textsuperscript{17} To be sure, these anarchists do not necessarily take action against states. Some anarchists avoid doing so for moral reasons, thinking that acting may cause more harm than good. Others avoid acting for prudential reasons. Yet all political anarchists believe that we should at least want the state to cease to exist in principle. Not so for philosophical anarchists.

Philosophical anarchism so defined has allowed for a diverse set of voices to be classified under its banner. William Godwin, arguably the first modern philosophical anarchist\textsuperscript{18}, provides what could be interpreted as an act utilitarian justification for philosophical anarchism.\textsuperscript{19} Under this interpretation of Godwin, utility

\textsuperscript{16} Reiman, \textit{Political Philosophy}, 47-52. At this point in Reiman's argument, the nature of these "bads" can be either harmful consequences or rights violations.

\textsuperscript{17} Oddly enough, I have yet to come across a term denoting the opposite of a philosophical anarchist. "Statist" is the first word that comes to mind, but that would seem to imply endorsement of something more than a nightwatchman state. This would be misleading, since it would obscure the fact that libertarians like Robert Nozick are also opposed to philosophical anarchists. See the first section of Nozick, \textit{Anarchy, State, and Utopia} for his defense of the legitimacy of the minimal state.

\textsuperscript{18} Philp, “William Godwin.” Godwin's status as the "first" philosophical anarchist is subject to debate. Specifically, his faith in the development of rationality sometimes makes him look like a \textit{political} anarchist who, like other political anarchists, thinks that we should actively work toward dismantling the state.

\textsuperscript{19} William Godwin’s “utilitarianism” is up for debate. For a strong argument in favor of that interpretation, see Clark, \textit{Philosophical Anarchism}, especially chapter 5. For an overview of the debate, see Philp, “William Godwin”. There are two points to be made here. First, if Godwin is a utilitarian, he’s an odd one. As Philp points out, he is (a) a perfectionist and (b) if a utilitarian, then one who, Philp says, “delivers utilitarianism from the more statist approaches of Bentham and later utilitarians.” Second, I am trying to describe various approaches to philosophical anarchism, not making an argument that any reasonable person would call Godwin a utilitarian.
will increase in step with our capacity to act rationally. He believes that governments tend to stymie the development of this capacity, with those political systems that concentrate authority in the hands of the few faring the worst. Godwin does accept that our current capacity to reason, immature (though improving) as it is, at least provisionally allows for the necessity of democratic, decentralized government, but this does not mean that we should presume that laws should be obeyed, all else equal. He believes that we should instead evaluate and choose to obey each law on a case-by-case basis since a presumption of obedience would be antithetical to the development of individual reason and judgment.  

To someone adopting the “Godwin as utilitarian” perspective, laws may serve as reminders or suggestions, but we should not assume that the utility gained from following any particular law will exceed that gained from flouting it.

While most philosophical anarchists past and present tend to be moralists, others have turned to philosophical anarchism as a consequence of rejecting morality. For instance, Max Stirner argues in *The Ego and His Own* that all fixed concepts, including moral concepts, risk alienating individuals from themselves and trapping them into self-imposed bondage. Stirner rejects essentialism in all its forms, arguing that "man" as an abstract being, as well as "the state" and morality itself are conceptual constructs which, while provisionally useful, come to dominate us. It is when we reify these concepts and use them to determine who we "really" are or what we ought

---

20 Clark, *Philosophical Anarchism*, 296. Godwin does however think that we should take the social disorder caused by law-breaking into account when deciding whether or not to comply.

21 Stirner, *The Ego and His Own*.

to do that we enslave ourselves to them to the detriment of our health and authenticity. Of course, Stirner would not consider inauthenticity a *moral* failing, since he is rejecting the supremacy of morality as a standard by which to judge oneself. So on what basis does he reject inauthenticity? Here it is helpful to conceptualize Stirner as a forerunner to Nietzsche and Freud insofar as he is giving us advice on how to avoid what we now might call neuroticism. Consequently, while Stirner would agree that the reified state is illegitimate, he would also reject the duty to dismantle the state, since that duty would imply adherence to a set of norms which are taken to be a higher form of moral authority. Hence Stirner, while clearly rejecting the legitimacy of the state, cannot be a political anarchist.

It is also possible to retain morality while jettisoning secularism. For instance, while some religious authorities exhort their members to actively establish a particular kind of polity, and others advise their members to accept whatever authority is in place, there are still others who reject the state's claims of legitimate authority while denying that followers have a duty to actively dismantle the state. Some Christian "anarchists" seem to fit this description. An adherent to this view would believe that a good Christian should be indifferent to the state since she should be focused on contemplating the Kingdom of Heaven. The presence or absence of the state would be an irrelevant consideration when determining how to act. As Mark Van Steenwyk points out, she might interpret *Romans 13* – where Paul advises us to "submit to governing authorities" – as a way of highlighting a more fundamental message on how

---

23 Carroll, *Crystal Palace*, 16.
to act toward oppressors in general. Citing Jacques Ellul's *Anarchism and Christianity*, Van Steenwyk suggests that "the passage thus counsels nonrevolution, but in so doing, by that very fact, it also teaches the intrinsic nonlegitimacy of institutions." In other words, the exhortation to obedience does not come from recognition of state legitimacy, but rather constitutes an elaboration of a more general moral posture that includes content on how to deal with illegitimate so-called "authorities." The upshot would be that the state is illegitimate but we do not have a duty to dismantle it.

II. Types of Philosophical Anarchism

II.1 Wolff and Simmons

For all its diversity, philosophical anarchism remained a theoretical backwater until it experienced a resurgence with the publication of Wolff's *In Defense of Anarchism* in 1970. Wolff modernized the doctrine by defending it on moral, non-utilitarian grounds and in doing so made it relevant to modern political philosophy.

Wolff ultimately concluded that states are illegitimate because there exists an irreconcilable tension between individual moral autonomy and legitimate state authority. His argument is (very) roughly Kantian and proceeds as follows. Individuals

---


26 I want to make two observations here. First, I understand that this interpretation of scripture is very controversial. Second, not all Christian anarchists are philosophical anarchists. Some are political anarchists who argue that Jesus’ call for non-violence would “ultimately imply a condemnation of the state for its theoretical and practical monopoly over the allegedly legitimate use of violence.” See Christoyannopoulos, “Christian Anarchist Critique,” 19-26. Others have claimed that state power is illegitimate because true legitimate authority rests with God, that God and only God can give legitimate commands that must be obeyed. Van Steenwyk, *That Holy Anarchist*, 52. I would argue (but not here) that people who hold this position are not anarchists of any type.

27 Alongside Wolff, Smith, Simmons, and Joseph Raz were among the most influential philosophical anarchists in the 1970s.
are autonomous because they have the ability to make reasoned judgments with respect to what they ought to do. This autonomy is what makes us moral agents. Our central duty as moral agents is to do our best to discern what is right and to act accordingly.\(^\text{28}\) Those who claim authority over others claim the right to be obeyed with respect to an issue or range of issues simply because they have issued a command. Authority is, for Wolff, “The right to command and, correlative, the right to be obeyed.”\(^\text{29}\) Wolff believes, in other words, that those who claim the right to be obeyed are asking us to either (1) stop making reasoned judgments except to determine that our relationship with the person issuing the command gives us a duty to obey, or (2) stop acting on the basis of the reasoned judgments we come to with respect to the content of the command. In other words: either we decide "Yeah, she has the right to tell me what to do here" and stop evaluating the morality of the action or we keep evaluating the morality of the action while obeying the other person even if this means acting contrary to our evaluation.\(^\text{30}\) But those who obey simply because they were issued a command are, Wolff believes, abandoning their autonomy.\(^\text{31}\) Autonomy can never coexist with authority, and thus the legitimate state, Wolff says, “must be consigned to the category of the round square, the married bachelor, and the


\(^{29}\) Ibid., 4.

\(^{30}\) Ibid., 9. Again, I should stress Wolff’s understanding of authority is reciprocal. As he says, “[Legitimate authority] is a matter of the right to command, and of the correlative obligation to obey the person who issues the command.” (emphasis his). I take this to be a standard view of authority, although I acknowledge that some have disputed this connection. See Edmundson, *Three Anarchical Fallacies*, chapter 2. Later I will argue that there will be cases where the state will be justified in creating and enforcing laws and that some philosophical anarchists will be justified in refusing to acknowledge the state’s claimed authority.

unsensed sense-datum.”

Commentators reacted swiftly and mostly negatively to Wolff's work. Indeed, Wolff has been assailed on multiple fronts. He has been accused of, among other things: not being a 'real' anarchist; offering an inadequate defense of a system of unanimous direct democracy that he concedes may address his concerns; being inconsistent in his claim that the just state is 'logically impossible'; being overly obsessed with moral autonomy at the expense of other virtues; confusing legal and moral obligations; and even irresponsibly advancing a dangerous thesis.

While most of these commentators rejected philosophical anarchism whole, a minority of influential thinkers, including M.B.E. Smith and Joseph Raz, instead defended modified versions of the philosophical anarchist thesis. It was the publication of A.J. Simmons' seminal *Moral Principles and Political Obligations*, however, which provided the foundations of what is currently the standard philosophical anarchist position. Simmons believes that he and Wolff agree qua philosophical anarchists that the state is illegitimate and that they do not have a duty to obey or oppose it. Yet Simmons goes on to claim that there are two types of

---

32 Ibid., 71.
37 Pritchard, "Wolff's Anarchism.”
38 Laszlo, "A Moralizing Note."
philosophical anarchist positions: a priori and a posteriori. A priori philosophical anarchists such as Wolff believe that “all possible states are morally illegitimate.” Simmons rejects this position and instead endorses a posteriori philosophical anarchism: “A posteriori anarchism, by contrast, maintains that while all existing states are illegitimate, this is not because it is impossible for there to be a legitimate state.”

Why the distinction if Wolff and Simmons agree that the number of existing legitimate states is exactly zero? The difference between the two is significant because only a posteriori philosophical anarchists can engage in a meaningful discussion with archists. Since Wolff believes that political legitimacy is necessarily impossible (since it would violate moral autonomy), he cannot engage with those who ask how it could best be brought about. Simmons has no such constraints. He could in principle be convinced by a powerful archist argument, and as such can specify the conditions that a state must satisfy to qualify as legitimate.

To help clarify matters, it helps to think of the debate in terms of three questions:

1. Could the state be legitimate?
2. If a state could be legitimate, what criteria must be met for a particular state to

---

39 The fundamental idea is in his 1979 work Moral Principles and Political Obligations, although to the best of my knowledge he first articulates the distinction in these terms in his 2000 paper, Philosophical Anarchism.
40 Simmons, “Philosophical Anarchism,” 105. In fact, and in contrast to Wolff's own remarks about married bachelors, Wolff does think that a hypothetical state would be legitimate if and only if it were governed by a unanimous direct democracy. See Wolff, In Defense of Anarchism, 21-27.
41 Simmons, “Philosophical Anarchism,” 105.
42 As I will argue in the following chapter, the disagreement between Wolff and Simmons turns on the question of consent.
be legitimate?

3. If a state were legitimate, what kinds of institutions, policies, or other considerations should we endorse?

Much democratic theory seems interested in answering the third question. Philosophical anarchists are interested in the first two. To use Simmons' terminology, *a priori* philosophical anarchists would insist that the answer to the first question is "no" and that the subsequent questions are therefore moot. *A posteriori* philosophical anarchists disagree. Their answer to the first question is "yes" and go on to say that there exists no state that satisfies the criteria laid out in their answer to the second question. Simmons has in effect successfully shifted the locus of debate between philosophical anarchists and archists from the first question to the second question.

Simmons presents three criteria that a theory of legitimation must jointly satisfy in order to confer moral legitimacy upon a given state. The criteria are: (1) accuracy – that the theory “identifies as politically bound those individuals falling within the scope of the principle(s) it uses”, (2) completeness – that it “identifies as bound all and only those who are so bound”, and (3) generality – that it “entail that most (or at least many) citizens in most (or many) states are politically bound.”

Crucially, there is no principled reason why these conditions could not be jointly satisfied, and by implication, Simmons must accept that some states may be

---

43 This is not to say that *a priori* philosophical anarchists would have nothing to say about how real states should be governed. They would, but these observations would have nothing to do with the state as a legitimate authority.

44 Simmons, *Moral Principles*, 85. For a discussion of this criterion and Simmons' standards of success more generally, see Klosko, *Political Obligations*, chapter 1.
legitimate. This sets up a winnable challenge for archists, and many have taken up the challenge with great enthusiasm.

The philosophical anarchist vs. archist debate has come to resemble a tennis volley taking place on a liberal court. It begins with the philosophical anarchists claiming that laws are coercive and that the burden of justification must lie with those who support them. Archists typically respond by drawing on one or more theories of obligation, including appeals to, among others: fair play, natural duty, utility, associational duties, and consent. Once an archist argument has been acknowledged, the philosophical anarchist strategy is to demonstrate its flaws, declare the strategy invalid, and wait for the next challenge. The fundamental idea is that until an acceptable theory of legitimation has been formulated – one that satisfies Simmons' three criteria, for instance – we should revert to the default position that the attempt to justify coercion has failed and that we do not, all else being equal, have a duty to obey the law.

Simmons, like Wolff, rejects existing theories of obligation but unlike Wolff seems to believe that consent theory holds the most promise to generate moral obligations. In fact, Simmons believes that a small number of people already have such obligations. Those who have explicitly taken an oath to obey a reasonably just state take on a moral obligation to obey the laws of that state. The problem, as

---

45 For an example of each, respectively, see: Klosko, *Principle of Fairness*; Rawls, *Theory of Justice*; Reiman, *Political Philosophy*; Horton, *Political Obligation*; and Jenkins, "Political Consent." Some have instead argued that the just state is not coercive – see for instance the second “fallacy” in Edmundson, *Three Anarchial Fallacies*, 73-125 – but I will avoid that line of debate here.

Simmons and others have pointed out, is that few people have made such an oath. As a result, consent theory fails to satisfy Simmons' generality criterion, although we could imagine a future case where this could change. A large number of people could consent, resulting in a legitimate state according to Simmons' rules.

II.2 “My” Country and the Political

Here I pause to return to the question of what I mean when I ask whether or not I have a duty to obey “my” country’s laws. What initially seemed to be a straightforward question, at least to me, is unfortunately quite problematic. Roughly, there are two types of answers here, and work by Tracy Strong shows why this is so.

In a recent article, Strong differentiates between political and economic enterprises. He defines an enterprise as political when “human beings are engaged in activities that constitute a simultaneous and common answer to the questions of 'what am I?' and 'what are we?' Note that 'we' is as much 'me' as the 'I'.” Economic activities are those which “at least in part constitutes an answer to the question of 'what do I have to do to get what I want given that other people also want it?'” Political discussion necessarily entails a negotiation (of sorts) about what it means to be a member of a shared group, e.g., "Americans." To use one of Strong's examples, if you responded to my claim that eavesdropping on private conversations is "un-American" by retorting that the country is threatened by terrorists, my saying that the Constitution forbids it would be a political response. My saying “That’s just how I

48 Ibid., 442.
feel” would not.49

In the second case, my answer constitutes a withdrawal from the conversation. In the first, however, my response continues a negotiation about what it means to be American, but not a negotiation in the utility-maximizing, bargaining sense of the term. Rather, the negotiation is an argument that “will thus necessarily involve humans in the stuff of power, authority, conflict, argument, discussion, compromise, leadership, institutions, and so forth.”50

Strong’s distinction helps us see that there are at least two ways of understanding what "my" means when I ask whether or not I have a duty to obey the laws of "my" country. One answer is that it references a shared relationship with others, a relationship which in part constitutes who I am. If I am American, this would mean that (1) being American is part of who I am and that (2) I cannot unilaterally declare what "being American" means. The status of "Americanness" will be contested with others who are also Americans, and this contestation will be messy and ongoing. Given that, the meaning of "my country's laws" in this political sense can be subdivided into two different possible statements. First, it could mean "the laws that currently exist in my country."51 If I am American, then it would mean the laws that are on the books in America, even if I disagree with them. The idea here is that even if I disagree with the laws, they are my laws (because they are our laws) and I am

49 Ibid.
50 Ibid., 441.
51 For the sake of this discussion, I am talking about citizens of a country who live in that country. Cases such as the status of expatriates are interesting, but only lead to unnecessary complications here.
responsible for them in the same way that Strong (sadly) acknowledged that George W. Bush was “his” president (because Bush was our president).\footnote{Strong, “Political Realm,” 450.} Second, "my country's laws" might mean "the laws that plausibly correspond to the essence of what it means to be American." Under this understanding, some laws that are on the books might not be "my country's laws" if they clearly violate very basic tenets of Americanness, or if they were brought about by usurpers who are not "American."

The other answer to the question of what “my” means need not be political in Strong’s sense of the word. “My country” could mean the set of institutions which offer a set of positive and negative incentives to promote certain behaviors. In other words, I can be quite alienated from my country, at least in the political sense. Strong would agree: “What this points to is the fact that human relations are entirely possible without being political.”\footnote{Ibid., 443.} In such a situation “… politics becomes more and more purely instrumentalized, more of a way of getting what you want.”\footnote{Ibid.} I should not here that “what you want” need not be purely self-regarding. I want, among other things, for my loved ones not to suffer, to feel safe when going about my business, and for my moral desires to become instantiated.

Unfortunately, philosophical anarchists are not always clear which of these interpretations they are adopting. As a result, they may fail to satisfy some archists even if they were to agree that they have a duty to obey the law or a particular class of laws. If after much reflection a former philosophical anarchist decided that he had a
duty to obey the laws of "his" country, he may do so without engaging in others politically, but rather because he believes that doing so would best instantiate his moral or subjective values.

This dissertation does not aim to argue that philosophical anarchists should approach the question in one way or another. If a philosophical anarchist decided to become an archist for reasons that concede nothing to Strong's idea of the political, I would consider those reasons successful. In other words, in this dissertation I am fine with "economic" interactions between philosophical anarchists and archists. That being said, it is also the case that philosophical anarchists may come to develop and be "captured" by political bonds. Different archist strategies will imply different types of relations. Utilitarian arguments tend to avoid making "political" demands (in Strong's sense), whereas arguments based on gratitude are more likely to argue that a political bond exists and should, therefore, give philosophical anarchists good reasons to obey the law.

The political / economic distinction is useful here for at least three reasons. First, it can help individuals engaged in the debate determine if the debate is worth having. Some philosophical anarchists – notably Wolff – seem intent on rejecting the political, and some archists would likely find an “economic” justification for the duty to obey the law unsatisfying. These types of people should either (1) recognize that they are talking past each other, and/or (2) discuss their disagreement regarding what framework – political or economic – is appropriate when discussing state legitimacy.

Second, the distinction helps illustrate an important difference between
philosophical anarchists and Reiman's "political" anarchists. Reiman correctly points out that the two camps differ in that the only the latter believe that there is some duty to undermine the state.\textsuperscript{55} Yet Reiman’s distinction may be too crude. His political anarchist’s thought process looks a lot like the non-political, atomized perspective that most philosophical anarchists seem to share. According to Reiman’s description, an anarchist could become a political anarchist from one day to the next when she decides that some value could be better actualized by eliminating the state.\textsuperscript{56} There is surprisingly little discussion of political (in Strong's sense of the term) solidarity that permeates many classical anarchist texts.

Finally, although there is a surprisingly big gap between the aims of philosophical and political anarchists, applying Strong’s framework to the political anarchist agenda can help illuminate a path that (might) allow philosophical anarchists to reject the state while accepting some political authority in Strong’s sense of the word. Political anarchists are typically very interested in developing groups with deep solidarity, strong group identification, and inclusive and nearly-infinitely contestable ways of deciding how resources and social relations are to be organized. They oppose the state’s practices and institutions, but not in the same way as typical protesters do.

While typical American protesters seem to take a line similar to what Rawls articulated in his \textit{On Civil Disobedience} – a protest on the very political grounds that a law or practice is in conflict with a higher-order value that is constitutive of what it

\begin{flushleft}
\textsuperscript{55} Reiman, \textit{Political Philosophy}, 48-52.
\textsuperscript{56} Ibid., 49.
\end{flushleft}
means to be American – political anarchists often deny that they are American at all. What they implicitly argue is that one can embrace the political while rejecting statist labels that they believe do not apply to them. Their stance is that we are enmeshed in relations with others, and that these relations contribute to a sense of "we" that is contestable and that imposes duties on us, but that these political relationships can and should exist absent the tether to what they believe to be currently harmful identities like "Americanness."

There is no principled reason that philosophical anarchists cannot, in the course of their discussion with archists, reach a position which concedes their participation in political relationships. For instance, perhaps when arguing that they do not have a gratitude-based obligation to obey the law (which I explore in the next chapter) they contrast their relationship to the state with their relationship with their local community. Were this the case, they would come to resemble political anarchists a great deal. This would be an intriguing development: a defense of political anarchism, complete with a respect for the duties stemming from solidarity, but without the Marxist pedigree that accompanies much anarchist thought.

III. Simmons vs. Simmons

Before moving forward with an investigation and critique of Simmons, a quick summary is now in order. Simmons agrees with Wolff that even relatively “good” states are illegitimate. Unlike Wolff, Simmons thinks that states could be legitimate
and that a successful general theory of legitimacy would satisfy three criteria: accuracy, completeness, and generality. Simmons believes that existing theories of legitimacy fail and that consent theory, which already morally binds some individuals to law abidance, has the greatest potential to eventually ground a general obligation to obey the law. That being said, however, I will argue that Simmons' views on consent ultimately undermine his remarks on the requirements of a valid general theory of obligation. Specifically, I aim to show that the generality criterion should be irrelevant to someone like Simmons, and that if this is so, we should call into question the earnestness of his back-and-forth with his archist interlocutors.

Simmons' philosophical anarchism is a direct result of his appropriation and modification of Locke's political and moral theory. In fact, Simmons is to Locke what Wolff is to Kant: a sympathetic follower who ultimately thinks that his predecessor's positive proposals go too far. In Wolff's case, the fallout is merely that some accuse him of holding a poor interpretation of his source text. Simmons, on the other hand, must face the added charge that his interpretation of Locke makes his discussion with archists irrelevant. I will demonstrate this in three steps. First, I will briefly describe Simmons' interpretation of, and agreement with, Locke. Next I will discuss Simmons' objections to Locke. Finally, I will discuss the upshot with respect to the previous section of this paper.

Simmons agrees with (his interpretation of) Locke's views on legitimacy, the state of nature, and the importance of consent. From Simmons' Lockean perspective, a

58 I discuss Wolff's Kantianism in the final chapter of my dissertation.
legitimate state is not necessarily legitimate with respect to those who inhabit its
territory, even to those who have lived in the state their whole lives. We are born in
the state of nature, and we can remain in this state while living under legitimate
governments.\(^{59}\) For Simmons via Locke, we only leave the state of nature when we
consent to form or join a political community.\(^{60}\) When we do so, we incur the
obligation to follow the laws of a government – so long as the government remains
within its jurisdictional bounds – out of an obligation to the political community that
gave rise to the government. If an individual does not consent to form or join the
community, the government is illegitimate (even if just or otherwise “good”) with
respect to that individual, even if it is legitimate with respect to that individual's
neighbors.

Simmons parts ways with (his interpretation of) Locke on the question of what
constitutes consent. Locke's problem, according to Simmons, is that he does not take
his thoughts to their natural conclusion and instead twists his philosophy to align it
with common intuitive conclusions about the nature of political obligation.\(^{61}\)
Specifically, Simmons thinks that Locke wants to say that legitimate states exist and is
willing to torture his theory in order to do so by arguing that residence constitutes tacit
consent. Simmons argues alongside Hume that residence cannot be the basis for

\(^{59}\) Simmons, *Edge of Anarchy*, 21-22.

\(^{60}\) Ibid., 59-60, 68-85, 171. This consent must be "actual consent," i.e. an actual act taken by a
particular person at a particular time. Simmons rejects hypothetical consent theories, which
"purport to derive our obligations to obey from the consent that would be given by some idealized
version of ourselves," claiming that such theories are merely natural duty theories with a different
name. See Wellman and Simmons, *Obey the Law*, 117.

\(^{61}\) Simmons, *Edge of Anarchy*, 194.
consent, and while he acknowledges the possibility that someone could consent tacitly, he seems very skeptical of the use of tacit consent to ground political legitimacy. The conditions for tacit consent to “count” are just too strenuous: “Failing to do something can only be a way of consenting when the inactivity is in response to a clear choice situation, only when inactivity is significant as indicating that a choice has been made (and, as we will see, not always even then.” Consent thus requires a clear signal, and residence certainly does not constitute that signal. Simmons believes that we live in a society where most people do not give clear consent and that if the conclusions we must draw from this run contrary to our archist intuitions, so much the worse for our intuitions.

There are a number of upshots here. First, Simmons believes we consent not to obey the government per se, but rather to form or join a political community which then establishes a government. To those who have consented to this, governments can become illegitimate without the dissolution of the political community which gave rise to the government. In such cases, we would, like natural duty theorists (which I discuss in Chapter 5), have a duty to replace the government with a legitimate one. This gives Simmons the flexibility to claim that consenters keep their basic political obligations even when their government no longer represents them. A government's failure to act appropriately does not mean that "anything goes" as far as political obligation is concerned: we are usually still part of a political community and must act

---

62 Ibid., 87.
63 Ibid., 221-222. Emphasis in the original text.
64 Ibid., 246.
accordingly.

But second, Simmons' "Lockean" philosophical anarchism seems at odds with his contention that a theory of legitimacy – one that would satisfy an earnest would-be philosophical anarchist – would succeed if it fulfilled his three criteria. Simmons says that consent theory fails because most people do not consent. Even if most people do consent, though, what is to be done with the non-consenters? Simmons seems quite sympathetic with those who, due to their unusual tastes, choose not to consent:

Thus, for instance, for those who genuinely value independence highly enough, any terms of membership offered by a modern state may be unreasonably harsh (since membership has a high disutility for such persons). To force such (admittedly unusual) individuals to choose between membership and emigration will still seem wrong, and agreements made will still seem morally suspect, even if the terms of membership are quite reasonable for the majority of residents.\(^{65}\)

It would be odd to say that such a person would be compelled to accept the legitimacy of the government which claims authority over her, even if most of her neighbors are consenters.

The trouble with Simmons' depiction of the legitimate state as described in Part II of this chapter is that it turns out to be merely descriptive. A state may be legitimate by broadly conforming to his three criteria, but this legitimacy only obtains with respect to consenters; non-consenters have no more reason to obey the law in a legitimate state than they do in one that is illegitimate.\(^{66}\) If this is right, then Simmons is being unfair when he engages archists. The archist who finds a country that satisfies Simmons' criteria would be told that a minority of people would still not be morally

\(^{65}\) Ibid., 242.

\(^{66}\) Although they may have duties to obey good laws, to oppose evil states, etc.
obligated to obey the law, and this would call into question the whole point of the exchange between the archist and philosophical anarchists. Archists want to establish the conditions for legitimacy because they want to know what it would take to say that all citizens have a duty to obey all relevant laws. If Simmons' definition of legitimacy does not allow this, then he and the archists, while seemingly engaging each other in a debate, are talking about two different things.

As far as I can see, there are three possible ways to try to overcome this problem without distorting Simmons’ premises so much that he would simply ignore the argument. The first is to tighten up Simmons' criteria so as to allow no "leakage." This would mean that a state is legitimate only if, among other things, it is legitimate with respect to all those over whom it claims authority. The second is to propose that either states have the moral authority to bring dissenters to within their moral jurisdiction or that philosophical anarchists have the duty to act as if the state is legitimate. The third is to retreat and argue against state legitimacy from a priori grounds. The first option constitutes the project of many archists and will be examined at length in subsequent chapters. The remainder of this chapter examines the second and third options.

IV. Paths Not Taken: Alternative Responses to Simmons

IV.1 Alternative Responses

67 When I presented a version of this chapter in my school's political theory workshop, an audience member suggested that a Marxist analysis would reject Simmons' liberal premises. The audience member was correct. I do not follow this kind of strategy because I want to keep the discussion within the bounds of the existing liberal discussion between philosophical anarchists and archists.
This section addresses possible responses to Simmons that I do not cover extensively in this dissertation. One set of responses acknowledges that a theory of legitimation won’t “capture” all citizens and proposes that (1) the government is entitled to make demands on all citizens nonetheless and (2) that dissenters should recognize and accept the government’s right to do so. Another type of response claims that the duties of dissenters change if enough people assent to the state’s claims of legitimacy. Note that these responses are not comprehensive. For instance, one could adopt a radical posture and say that we should abandon Simmons and try to rehabilitate Wolff. While I find such efforts interesting, I do not pursue that line of inquiry here, although I return to this in the final chapter.

The first option is to say that when enough people are “captured” by a theory of legitimation, the theory can claim that those who slip through the cracks can nevertheless be justifiably coerced by the state and that they should recognize this coercion as legitimate. Robert Nozick has argued something like this in Part I of his *Anarchy, State, and Utopia.* In his hypothetical model, citizens in a roughly Lockean state of nature establish and freely contract with protection firms. These firms eventually merge, ultimately forming one large protection agency. Once a large enough proportion of citizens sign up with the agency, Nozick believes that the firm, by virtue of its obligation to supply its members with consistent and stable rules of order, will be entitled to coerce the remaining non-subscribers to obey its "laws." The

---

68 Tellingly, Part I is called "State-of-Nature Theory, or How to Back into a State without Really Trying."
firm is then bound to extend its protection to those it coerces, and the state is (hypothetically) formed.70

“Coercive moral authority" in this sense thus means that (1) the state is legitimate with respect to most people, (2) the state is not legitimate with respect to a minority of dissenters, (3) the state will coerce the dissenters, and (4) the dissenters, while not captured by the theory of legitimation that has captured the majority, acknowledge the state's right to coerce them. This means that there is no pretense that the dissenters are captured in the same way as is the majority, and thus no veneer of a deeply unified polity. What's more, in "coercive moral authority," the majority can make and implement decisions with the understanding that the government apparatus does not represent the groups in the same way.

These types of solutions raise at least two questions. First: how are the majority to govern given the acknowledgment that there are essentially two classes of citizens? One option would be to act as if the dissenters were the same type of citizens as those who unproblematically accept the state’s legitimacy. Another would be to legislate *as though* this were the case, but only as a device to create laws that apply equally to all. A third option would be to create and implement legislation with a full appreciation that a bloc of citizens is being coerced.

Second, what does it mean for the dissenters to “acknowledge” the state’s right to coerce them? If we use Wolff’s definition of authority – the right to command with

---

70 Ibid., 54-57; 88-118. To be clear, Nozick uses this story as a device to argue for a state under Lockean conditions without the type of consent that Locke finds necessary, and not as an attempt to describe how states actually came into existence.
the correlative duty to obey – it would appear to mean that philosophical anarchists would conclude that their duties would change once a certain proportion of people were captured by a theory of legitimacy. Such a strategy would not convince most philosophical anarchists, who would demand to know why there should be a connection between a majority’s falling under an archist theory of state legitimacy and their own duties.

One possible response is to decouple the right to command from the duty to obey. 71 I will not discuss this option at length, but will instead point out that this is probably not what most archists have in mind. Accepting that philosophical anarchists do not necessarily have a duty to obey concedes more than the archists are willing to give. I suspect that most philosophical anarchists would be happy with this reply since it vindicates their central claim.

Archists could instead try to create a link between widespread acknowledgment of state legitimacy and a duty to obey the law. Peter Singer’s strategy in Democracy and Disobedience could support an argument of this sort, although that is not exactly what he says. Singer’s model involves two conceptual steps. First, we have a duty to participate in fair political procedures, specifically voting. Second, our participation in these procedures gives others a legitimate expectation that we will abide by the outcome generated by these procedures, all else being equal. 72 Singer calls this "quasi-consent," which he likens to estoppel. As he says, "The effect of this doctrine is to prevent someone denying something which, by

71 Greenawalt, Law and Morality, 47-61.
72 Singer, Democracy and Disobedience, 59.
his voluntary behaviour, he led another reasonably to believe. In estoppel, as in what I have called quasi-consent, the real intention is irrelevant.” The natural duty to participate in fair procedures gets the people in the game, and playing the game causes them to incur obligations since others develop reasonable inferences about the meaning of the action.

The problem with Singer's model is that only those who actually fulfill their obligations seem to incur the additional burdens accrued from quasi-consenting by virtue of their actions. But most political action is not well-publicized. I am a political scientist, and even I can’t accurately gauge the political involvement of many of my acquaintances. A Singerian approach could hold that as more and more people participate, this model develops a “critical mass” criterion much like Nozick's: the more people endorse and abide by a given procedure, the more they develop expectations that all participants will treat the procedure as legitimate. Unless someone were to advertise their non-participation aggressively, others could reasonably expect them to see the government as legitimate.

Singer’s model is an interesting one, and I will return to it briefly in Chapter 4, where I discuss the connection between voting and consent. For my present purposes, however, I want to make three observations. First, strictly speaking, Singer’s scheme neither (1) accepts that it does not capture everyone, nor (2) tell a story about how those who are not (immediately) captured nevertheless incur obligations. His is a natural duty theory, and we all have the duty to participate. If his theory is right, then

73 Ibid., 51.
non-participants are the ones at fault. Singer shouldn’t have to tell a story binding those who avoid one duty by failing to fulfill another.

Second, as I will discuss later, philosophical anarchists could push back at my extension of Singer's argument by claiming that, assuming they have the right not to participate, then they would suffer undue burdens if they had to make a show of this non-participation. The amount of effort it would take to combat others' (reasonable) expectations might be too much to demand.

Finally, I want to highlight Singer's two-step model because I will propose something quite similar in the latter chapters of this dissertation. Singer suggests we have a duty to participate. I instead will propose that we have a duty to deliberate. The main difference will be that whereas Singer's model entails that someone has a duty to obey the law once they participate in his process, mine does not. The legitimating power of deliberation will, I will argue, depend on how it takes place.

**IV.2 Upshot**

In sum, if my observations hold, then both philosophical anarchists and archists must give some ground. Philosophical anarchists must come clean and clarify what it would mean for a theory of obligation to capture most – but not all – individuals to whom it ostensibly applies. What's more, if they acknowledge that it is ever possible to justifiably coerce innocent people, then they must recognize and squarely address the possibility that archists are well within their moral rights to coerc them and act as if they – the philosophical anarchists – have a duty to obey the
law. Philosophical anarchists could retreat to the belief that they do not acknowledge the authority of law, but, if they also believe that others may justifiably coerce them, this brand of philosophical anarchism becomes very, very thin.

The archists, on the other hand, should acknowledge that no existing theory of obligation has decisively settled the problem of political obligation. Some archists may say that this is because others have simply not come around to recognizing the virtues of the "correct" theory. These archists will batten down the hatches and keep on arguing for their pet theory. I examine some of these schemes in the next few chapters of this dissertation. Others, however, may accept that no theory of obligation that decisively captures all the individuals that it targets is forthcoming in the immediate future. These archists would do well to shelve attempts at finding find such a theory and content themselves with defending a model which would allow them to justifiably coerce recalcitrant but earnest philosophical anarchists. However, this may involve a reconceptualization of the body politic as all participants come to terms with the idea that *E pluribus unum* may not apply even in just states.

**V. Conclusion**

This chapter has advanced a number of claims which, if correct, imply that we must reconceptualize the root debate about state legitimacy. The dominant discussion between archists and a posteriori philosophical anarchists has taken place under unclear terms. On the one hand, philosophical anarchists like Simmons seem to be saying that an acceptable theory of legitimation needs to capture most, but not all
those it tries to capture. On the other, they offer no clear reason why a non-consenter would have a duty to obey the law even under such an "acceptable" theory.

Ultimately, liberal archists can advance the debate in one of two ways. The first is to take the position that an acceptable theory need not capture everyone it seeks to capture. In that case, we would need to come up with a theory of justifiable coercion. The second way of moving the debate forward is to redouble efforts to find a theory that goes beyond being “good enough” and actually does capture everyone it tries to capture. I address this latter strategy in the bulk of this dissertation.

The remainder of this dissertation is divided into two parts. In the first, I trace some major archist responses to the philosophical anarchists and then describe the philosophical anarchists' response, focusing on those which I believe allow for novel lines of inquiry. I examine gratitude-based arguments in Chapter 2. In Chapter 3, I discuss what I take to be the most promising version of fair play theory. Chapter 4 focuses on a particular variant of consent theory. Rather than rehash well-established theories of explicit consent or tacit consent based on residence, I consider the argument that the act of voting constitutes consent. One interesting feature of this type of consent theory is that the ostensible act of consent often comes about as a result of the discharging of a perceived duty, that is, the duty to vote. This has the same form as the tentative proposal that I will make in subsequent chapters.

Chapter 5 marks a transition to the second part of the dissertation and consists of two main parts. The first, shorter part summarily examines received archist arguments based on natural duty. The second discusses the bases of natural duty as I
understand them. I ultimately propose that natural duty theories of obligation have the
best chance of convincing philosophical anarchists, but that they fail to establish
political legitimacy as per Simmons’ three criteria. That said, I suggest that while
natural duty may not provide the direct solution that archists hope for, a duty to try to
deliberate might ultimately do the job.

I discuss deliberation in Chapter 6, tracing what I take to be the theoretical
evolution of deliberative democracy from a set of standards aiming to reject irrelevant,
non-reasoned discussion to a broad set of principles and practices that aim to ensure
the inclusion of all relevant voices. I ultimately argue that we have a duty to try to
deliberate and that deliberation, if done properly, can lead to political legitimacy. This
is easier said than done. For deliberation to legitimize, it must be done respectfully.
While all theories of deliberative democracy advocate respect, they tend to either insist
on behavioral standards that exclude the voices of the oppressed or do not show what
respectful behavior looks like.

I try to draw out some complexities of deliberation in the next two chapters. In
Chapter 7 I discuss the importance of etiquette in deliberation. Rules of etiquette in a
pluralistic society are far from clear, especially when individuals cross social
boundaries and interact with those they would normally want to avoid. Not only can
etiquette be violated – causing real offense – but it can also reify unjust patterns of
domination. In such cases, some deliberators may believe – justifiably, in my view –
that the deliberation is disrespectful and thus fails to legitimize. I then discuss apology
in Chapter 8. The same considerations apply here: apologies will be necessary for any
serious deliberation, and handling them in a respectful manner is far from trivial. Ultimately, while I believe that we may have a duty to deliberate and abide by the results of respectful deliberation, it seems that many will have good reason to feel that deliberation was disrespectful. Those people could then argue that they satisfied their natural duty to (try to) deliberate without taking on the duty of abiding by the results of the deliberative process.

I began this project with the hope of providing a clear answer to the likes of Simmons. The result of my work in Chapters 2-8 is a far more hesitant answer, or set of answers, that may disappoint some readers. With this in mind, I included a final chapter that switches gears dramatically. The bulk of my work engages and attempts to further a debate between a posteriori philosophical anarchists and archists. A more radical approach would be to try to rehabilitate the a priori philosophical anarchist position. While such a project would demand an entirely new thesis, in Chapter 9 I provide an argument that one of the main objections against Wolff’s philosophical anarchism is not as strong as it seems. Philosophical anarchists who are dissatisfied with Simmons’ more conciliatory stance would do well to continue the project of responding to Wolff’s critics.
Chapter 2: Gratitude

I. Introduction

Gratitude theory is currently the least respected serious theory of obligation. As John Horton points out, “Unfortunately, most references to gratitude as an explanation of political obligation are extremely brief and underdeveloped; and though it has recently been the subject of renewed interest this has been mostly of a critical sort.”\(^7\) That said, the theory goes back at least as far as Plato's *Crito*, and a few commentators – notably A.D. Woozley and, unsurprisingly, A.J. Simmons, have taken it seriously. This section relies heavily on their work.

But why not let sleeping dogs lie? Why devote a chapter to an unpopular theory? I have at least four answers to this. First and most prosaically, I do so for the sake of completeness. Second, I believe that although ultimately untenable, gratitude theory deserves more respect than it gets. My impression is that interest in one theory of obligation over another is partly due to circumstance and fashion. If I am right about this, it is by no means a waste of time to prepare a response in anticipation of gratitude theory's eventual return to prominence. Third, my discussion will hopefully add new distinctions both in favor and against such theories. In fact, in the course of writing this chapter, I had moments where I began to persuade myself of gratitude theory's viability. Finally, this discussion will allow me to introduce some theoretical constructs that I will use in subsequent chapters.

This chapter is divided into two parts. The first part presents what I take to be

\(^7\) Horton, *Political Obligation*, 101.
the strongest arguments in favor of gratitude theory. In so doing I will re-organize and in some cases add to the existing literature. The second part argues against gratitude theory and finds that it should prove unpersuasive to philosophical anarchists.

II. Description

At their core, all gratitude theories rely on the intuition that we have some duty of gratitude to those who confer benefits on us, so long as we should or do recognize these “benefits” as in fact beneficial and that they are not being conferred as part of a mutually beneficial contractual obligation. These duties can be understood in a number of ways, and so this subsection investigates the answers to two questions that I believe any gratitude theorist must answer: (1) what kind of relationship do the parties have?, and (2) what do duties of gratitude consist of? I will take each of these in turn. I will then outline a number of difficulties that gratitude theorists must try to overcome.

II.1. What kind of relationship? Filial vs. Transactional Gratitude

II.1.a Filial Gratitude

Gratitude theorists tend to argue by analogy, and the most obvious analogy here is the comparison of the relationship between citizen and state with that of child and parent. Horton (among others) argues that this is what Plato had in mind in the Crito, saying that Plato's position is that “In both cases gratitude is merited because of ...

75 I should note straight away that duties of gratitude do not necessarily include feeling grateful, although there may be good reasons, and possibly a separate duty, to cultivate a disposition that would lead one to feel grateful at appropriate moments. I discuss this at greater length in section II.2
the succour and support which they have provided.” And while Horton immediately dismisses the analogy as implausible, Simmons takes it seriously enough to explore its appeal, although even here his bias toward contractual thinking ultimately shines through.

That children owe their parents a debt of gratitude is not immediately implausible. After all, *good* parents do provide “succour and support” to their children, and while the parents (hopefully) enjoy this relationship, they also provide far more than they receive. As Woozley says, “... the children of a normal family are better off for what their parents have done; furthermore, much of what the parents did would have been done over a period during which the children could not have done it for themselves”. The parents may be better off for having had children, but the relationship is asymmetrical in at least three important ways.

First, as much as some parents love their children, what the children provide in many cases does not seem to reach a level similar to the “support and succor” that they have received, at least in the US and many parts of Western Europe. Hence the commonplace notion that children should not expect personal gratitude from their parents, but should instead expect gratitude from their own children. Second, the

---

76 Horton, *Political Obligation*, 100. I should point out here that the literature says little about what to do when the parent turns out to be mediocre. For the sake of this paper, I assume (unless I say otherwise) that parent is competent and well-meaning. I should also point out that Horton’s portrayal of Plato (via Socrates) may be incomplete, since Socrates also seems to claim that the Laws have shaped him and that they are owed obedience for that reason. It is an argument which, at least in the way that Socrates seems to present it, places the Laws in a relationship of parental authority with respect to citizens. 


78 Parents may be grateful for their children’s existence, but, as I will discuss later, this is different from being grateful to their children and thus incurring debts of gratitude.
nature of the asymmetry becomes clear when we imagine a loveless relationship between parent and child. As long as the parent does an adequate enough job for the child to grow into a functioning adult, she can argue with a straight face that the child owes her a debt of gratitude. It would border on nonsensical for the child to make a reciprocal claim. This points to the third difference: ideally, both child and parent feel grateful for the mere existence of the other, but the child also should – according to gratitude theorists – be grateful for what the parent actually does.79

What's more, Woozley's description provide an answer to those who argue that children are in no position to refuse the benefits their parents hopefully provide.80 Children are so vulnerable that we can reasonably assume that a lack of acceptance – indeed, it may be possible that they are too incompetent to "accept" goods in a way that would satisfy a rigorous political theorist – should be overridden on the grounds of hypothetical consent. What is important to the hypothetical consent theorist “is not whether consent is actually given but the basis upon which one ought to consent.”81 These grounds are usually understood to be moral, but hypothetical consent can also be brought to bear when we believe that someone is unable to act in accordance with their best interests. Children lack the competence to make responsible decisions, so we feel comfortable saying that they ought to want what good parents would provide,

79 Note that having children may make people worse off. Media outlets have seized on a recent study in the journal Demography and published reports suggesting that parenthood leads to a decrease in happiness. For example, see Cha, "Parenthood is Worse." Further research slated to appear in the American Journal of Sociology has validated this. For a brief by the authors, see Glass, “CCF Brief.”

80 Fair play theorists will not have this luxury, as I will discuss extensively in the next chapter.

81 Horton, Political Obligation, 84. For the seminal works on hypothetical consent, see Pitkin, “Obligation and Consent I” and “Obligation and Consent II.”
even if we do not blame them for not wanting – or believing that they do not want – those things. In other words: children need the benefits they receive, and if they do not believe that this is the case, we feel comfortable saying that they are just mistaken. Children who refuse care are taking a radically self-destructive stance, one that they will predictably regret having been allowed to take. They must accept benefits that we provide and must therefore, the argument goes, incur duties of gratitude, whether they realize it or not.82

Still, some might argue that contracts (implicit or explicit) are what generate duties and filial duties do not exist because children are too incompetent to have any duties at all. Call this the naive contractarian approach. It states that children who are not old enough to enter into contracts cannot be bound by them and thus have no duties. A modified and more palatable version of this might be that those who are unable to enter contracts cannot incur duties.

Yet even Simmons, consent theorist extraordinaire, does not think that all obligations arise from contracts. He sensibly believes that some duties, limited though they may initially be, come about as we gradually mature.83 In other words, we acquire these non-contractarian duties as we grow old enough to understand and act according to them. I hold my six-year-old niece to a different moral standard than I do her two-year-old sister, and I will hold her to a still different standard as she grows older. We

82 Children are odd cases of hypothetical consent if we assume that they are never able to give consent. Most situations where hypothetical consent is involved are cases were an adult could in principle give consent but does not do so due to incapacitation, mistaken beliefs, or an ill will.
83 Simmons, “Obligations of Citizens,” 51-55. This makes sense given his brand of Lockeanism. We are endowed with prepolitical rights, some of which may be waived when entering into certain types of contract.
frequently and coherently tell children that they ought to be ashamed for having stolen and broken her classmate's toy, and it makes sense to tell a high school student that he has a duty not to cheat on his exams (or romantic partner, for that matter). At the very least, Simmons points out, children would owe their parents those duties that they owe to others.\textsuperscript{84} And if children hold some duties toward their peers, it is also at least reasonable to consider the possibility that children hold additional duties to their parents, either on the grounds of the sacrifices that the parents made when the children were in need or out of familial ties of love.\textsuperscript{85} The upshot: filial gratitude should not be laughed out of court.

Before discussing non-filial gratitude, it is worth highlighting a further distinction made by A.J. Woozley, as doing so will prevent us from taking a wrong turn. In his book-length analysis of political obligation in the \textit{Crito}, Woozley distinguishes between what he calls the \textit{filial gratitude} argument and the superficially similar \textit{parental authority} argument. On the latter: “The argument that laws, like parents, are owed obedience in return for the benefits they have conferred on those subject to them merges in 50e into the argument that laws, like parents \textit{and masters}, are owed obedience because of their \textit{authority} relationship to the subjects.”\textsuperscript{86} In other words, parents have authority for reasons that may not have to do with gratitude and they should be obeyed for those reasons.

Clearly, this is not the type of gratitude that I am discussing here. In the filial

\begin{itemize}
\item \textsuperscript{84} Ibid., 51.
\item \textsuperscript{85} Ibid.
\item \textsuperscript{86} Woozley, \textit{Law and Obedience}, 70.
\end{itemize}
gratitude account, the kind that I have discussed and will critique below, the authority of parents with respect to their children derives at least in part from duties stemming from gratitude. Gratitude, in other words, would account for a subset of the duties that children owe their parents. To the degree that it does so, it is responsible for whatever authority parents have, at least from their children's perspective. The parental authority position, on the other hand, presumes obedience even before the question of gratitude comes up. The parental authority argument seems so antithetical to a liberal framework that I simply won't deal with it in this paper. The filial gratitude, on the other hand, does not presume the authority of parents, the state, or "masters," and is relevant enough to be rebuked.

II.1.b Transactional Gratitude

The more we look at the "filial gratitude" interpretation, the more likely we are to decide that the analogy is superfluous. If we owe a duty because we are receiving benefits, why invoke family relationships at all except as a superficial example – but not a model – of such duties? Perhaps the comparison to the family confuses more than it clarifies and should be jettisoned altogether. This appears to be how Simmons interprets gratitude theory.87 From this perspective, gratitude is not relational, but transactional. Duties of gratitude can and should be understood strictly in terms of costs incurred by the provider and benefits conferred to the recipient. The fact that parents and children have some romanticized special bond is irrelevant here; the only

87 See especially Simmons, Moral Principles, 183.
thing that makes parents special from a transactional perspective of gratitude is they are uniquely placed to provide the benefits that would generate duties of gratitude on their children.

Transactional gratitude should not be understood as filial gratitude minus hokey appeals to sentiment. Consider the following instructive assertion by Simmons: “The obvious point that bears on the moral significance of parental provision of benefits is that it is often the parents' duty to provide them. The care, attention, and healthy environment that a good parent gives his children can hardly be compared with a gift from some unrelated individual.”

There are numerous points worth considering here. First, any duty owed is due to benefits provided. Second, these benefits need not be strictly material; care and attention also count. Third, even Simmons acknowledges that there is something special about family. Someone favoring the transactional interpretation of gratitude may nonetheless insist that the "specialness" of parents consists of their unique position to provide benefits to their members. And fourth, and most importantly, substituting "strangers" for "children" will have significant ramifications. I will discuss these in section III of this chapter, but to foreshadow: filial gratitude will have an easier time navigating the distinction between accepting and receiving benefits since it has a more plausible recourse to hypothetical consent.

Is this transactional gratitude necessarily pre-political? No, not necessarily. It is true that transactional gratitude is more appealing to the kind of liberal philosophical

anarchist that is the subject of this dissertation, since it allows us to conceptualize duties in terms of isolated actions taking place between two atomized individuals. And it is true that we, or at least I, typically imagine gratitude occurring between members of a political group – and that is based on actions taken with those political identities in mind – as quickly turning into the kind of hierarchical gratitude structure displayed by filial gratitude. Yet it is nevertheless possible to imagine an anarchist society that is properly political, featuring an egalitarian ethos but where duties of gratitude emerge not with respect to the political community as a whole, but to the subcommunities with which an individual is interacting for the present. In fact, one of the hallmarks of these idealized societies is the ability for a member of a subcommunity to move to another subcommunity if she feels as though she can enter into a better relationship with the latter. 89 These relationships will typically persist as long as the subcommunity members feel as though they have a duty to contribute to their neighbors and co-workers because others are helping them as well.

These two types of relationships – parental and transactional – will form the basis for the remainder of the discussion, but there exist other, less prominent understandings of relationships that produce duties of gratitude. A.D. Walker, for instance, argues that citizens owe gratitude as a result of ties of "goodwill and respect". 90 And as Chaim Gans points out, others, notably Philip Soper, appear to

89 This is a feature in much of the classical anarchist literature. Incidentally, one need not be a Marxist or defend an offshoot of Marxism to endorse this utopian vision. William Godwin, an almost cartoonishly staunch individualist, supports this view. For a comprehensive account of Godwin’s decentralized liberal anarchism, see Clark, *Philosophical Anarchism*, chapters 7, 9–10.

90 Walker, “Political Obligation,” 210-211.
argue that we owe a duty of gratitude toward those individuals who enforce the laws.\textsuperscript{91}

I will leave these positions aside. Horton says that positions based on respect for others, if valid, are simply not theories of gratitude, and I agree..\textsuperscript{92} Respect-based accounts of obligation may be grounded on fairness, natural law, or reciprocity, but not gratitude. Soper's argument (as Gans presents it) is vulnerable to the same critiques that apply to more prominent theories of gratitude that I explain below, notably the problems attached to the intentionality of the "benefactors."\textsuperscript{93}

II.2 Psychological and Moral Gratitude

The last distinction I want to make in this section points to a difficulty in answering of what gratitude actually consists. Gratitude is colloquially thought of as a psychological state. On this interpretation, we are grateful when we feel a certain way toward another. This feeling is difficult to describe accurately, but it probably contains the following traits. First, it is a pleasurable sensation, although it may be bittersweet. Second, it involves an awareness of the good others have done – or have attempted to do – on our behalf.\textsuperscript{94} Third, there must be some recognition that the benefactor did not have to perform the action(s) which produced the feeling.

For example, my students are unlikely to be grateful for my services if they

\textsuperscript{91} Gans, \textit{Philosophical Anarchism}, 45.
\textsuperscript{92} Horton, \textit{Political Obligation}, 102.
\textsuperscript{93} Gans cites Soper's "The Obligation to Obey the Law," published in 1989. Soper's more recent work, notably \textit{The Ethics of Deference}, continue to emphasize our relationship with those individuals who enforce the law but does so on the basis of fair play and promise-keeping, not on gratitude.
\textsuperscript{94} As I shall indicate in the section describing the challenges facing gratitude theory, there can be disagreement as to what agent is the proper focus of gratitude in some instances.
believe that I am doing my job only to get paid. If someone points out that I'm showing up and doing work that benefits them, they are likely to say "Well, he has to do his job. Besides, he knows he'll get fired if he doesn't." They might feel grateful that they have me as an instructor, but they are not grateful to me for my teaching the class and grading their assignments. We usually do not feel gratitude when others fulfill basic or contractual obligations. If my students feel grateful to me, it is because they believe that I somehow went "above and beyond" my required duties.

This lack of gratitude when considering someone who does what he "has" to do also applies to moral duties. We usually do not feel grateful to others for being decent people who discharge their basic duties, even if they have no clear material incentive to do so. Again, the question is whether or not people go “above and beyond” and perform supererogatory actions. Someone might be grateful if I post “Happy birthday” on their Facebook wall after having been reminded to do so by Facebook and after having been provided a convenient mechanism to do so with absolute minimum effort on my part, but we would not think poorly of the person's character if they felt little or nothing in response to my perfunctory birthday wishes. Contrast this with how we would feel if the person received a well-thought out handwritten letter from a friend that described in heartfelt detail the evolution of their relationship over the years. Dismissing this gesture would be more worrisome since

---

95 Being a decent person is similar to how comedian Louis CK describes being 40 years old: "Nobody is like 'I helped a forty years old guy today and it felt really good to do something for him.' Nobody spends their holidays delivering hot meals to forty-year-olds. And you're not young enough for anyone to ever be proud of you or impressed. They're like 'Yeah, do your job. […] That's what you're supposed to do.'" C. K., *Chewed Up.*
the letter writer went above and beyond her basic duties of friendship. We would at least feel as though the letter-writer is (and should be) far more likely to evoke feelings of gratitude.

These proper” feelings of gratitude are to a large extent learned – we find out that some behaviors are to be expected and are not subject to feelings of gratitude.\textsuperscript{96} Psychological gratitude is therefore not just a subjective state, but also an indicator of whether or not a person has been properly socialized, at least from the perspective of her community. It is this socialization that connects the subjective feeling of gratitude with \textit{moral} gratitude. The socialization is how we tell members, and especially children, that they ought to feel indebted in specific types of situations and that this feeling is an indication that they have incurred moral duties toward their benefactor.\textsuperscript{97}

\textit{Moral gratitude} is what we might call the set of duties that exist in situations where we would normally expect psychological gratitude. To the extent that we are properly socialized, psychological gratitude and moral gratitude should coincide.\textsuperscript{98} Yet this conceptual separateness is important; there are many cases where they do \textit{not} coincide. Consider a case where we tell a child that she ought to be ashamed for being

\textsuperscript{96} This applies both in terms of how we expect others to act as well as what we learn we are entitled to by virtue of our status as human beings, our social position, our particular position within a specific interaction, or some combination thereof.

\textsuperscript{97} The method of this socialization will vary, but I believe that philosophical anarchists would only be concerned with psychological gratitude in two respects, at least as far as their debate with archists is concerned. First, they are concerned with moral rightness, not sentiment per se. How one feels is not relevant, prima facie, when judging the worth of one’s action. Second, they would surely object to socialization that seriously threatened an individual’s autonomy – that is, their ability to make judgments and act on those judgments. However, I have not come across recent philosophical anarchist texts that worry about such matters. I therefore assume that the question, while fruitful in its own right, is out of place here.

\textsuperscript{98} Simmons makes a similar distinction, differentiating between gratitude as a psychological state and gratitude as an outward performance. See Simmons, \textit{Moral Principles}, 166-167.
so ungrateful. Moral education is taking place, but (1) we may successfully educate her into believing an incorrect statement, (2) we may fail to educate her, leaving us with a small human who does not feel the call of duty when we think she should, or (3) we may educate her "too well", leading her to overgeneralize the lesson and feel shame in inappropriate circumstances. She may come to feel as though she has duties when she, in fact, has none. This kind of "overlearning" can be hard to detect and, even if noticed, is frequently not corrected. Gratitude theory, besieged as it is, would be extremely fragile if it stated that we only had duties of gratitude when we actually felt grateful, since our feelings can easily lead us astray. What ultimately should count for our purposes is determining whether or not moral gratitude applies to questions of legitimacy.

III. Critiques

III.1 Against filial gratitude

This section presents the case against. First, and most superficially, several authors believe that the analogy is simply too stretched to be taken seriously. Surely many would agree with Gans' colorful assessment of filial duty position: “The Socratic position is the weakest one. It seems metaphysically undesirable to personify laws (what metaphysical consideration would justify a conception of laws as the same type of entities that humans are?), while the merciful character Socrates assigns this metaphysical embarrassment presupposes just what it sets out to prove.”\(^99\) Or as

Simmons puts it more circumspectly: “But for this assumption [that the analogy holds] to be justified, more points of similarity will be required to bolster the alleged analogy than the mere fact that benefits are provided in both cases.”\textsuperscript{100} This is not to say that the analogy never made sense. Perhaps it was quite compelling at certain times. Perhaps some societies today would find it persuasive. Yet it does not do much to add to a discussion among people who, for better or for worse, are more or less committed to some sort of individual-centric liberalism.

The analogy is made even more suspect by the increasingly technocratic and bewildering functioning of the state. As Woozley points out, “The kind of feeling which it is natural for a person to have for his parents it is hard to have towards laws, especially under modern conditions where they usually impinge on the subject's consciousness only when they frustrate him, fail to give him what he thinks they should, or bewilder him by their obscurity.”\textsuperscript{101} If the state is a parent, it is a very strange one. Gone are the days where a citizen can meaningfully believe that he is addressing the state – a loss which incidentally has fueled the recent interest in deliberative democracy, which I will discuss at length later. For most citizens, interactions with the state are at best mildly stressful events and at worst something approaching Kafka-esque bewildering ordeals. Prolonged direct interactions with the state tend to mirror the famous statement about war: mostly periods of boredom punctuated with bouts of sheer terror. If this is analogous to a parent-child relationship, we might excuse the child from her ostensible duties of gratitude.

\textsuperscript{100} Simmons, Moral Principles, 162.
\textsuperscript{101} Woozley, Law and Obedience, 68.
But what if the state acted like a decent parent? In such cases, the filial gratitude argument might still fail. Some theorists have argued that we might not owe duties of gratitude towards our parents after all. Simmons once again takes the lead on this point in *The Obligation of Citizens*. His position can be summarized as follows. First, some children are simply too young to have obligations of any kind. As children mature, they come to acquire general duties as they gain the ability to understand and discharge these duties. These are natural duties and not resultant from a special relationship with their parents.

If there is a set of duties that emerges out of the parent-child relationship, Simmons believes that it is the parents who have special duties, not the other way around. He gives two reasons for this. First and most obviously, some parents have accepted certain responsibilities that come with willful procreation. Second, even in cases where there is no willful procreation, parents have the right not to be interfered with (to some extent) when discharging their parental duties, although this right can be revoked if society deems them incompetent. In exchange for this right of non-interference, the parents agree to provide for their children in special ways. In other words, discharging their special duties allows parents to raise their children as they see fit, within reason. Again, it is the parent, not the child, that has the special duty here.

All in all, the analogy is unlikely to convince philosophical anarchists of much.

---

102 Simmons, *Obligations of Citizens*, 51.
103 Ibid., 52.
104 Simmons makes the rhetorically wise decision to avoid discussing accidental pregnancies, although I believe he would say that except in exceptional circumstances, “negligent” sexual activity is like drunk driving: lack of willful intent does not allow one to escape one's duties.
But as I indicated earlier, one need not rely on the analogy to be a gratitude theorist; one can also be a transactional gratitude theorist. The remainder of this section deals with this variant of gratitude theory, although some of the arguments also apply to the filial gratitude account.

### III.2 Intentionality

Duties of gratitude, if they exist, only come about when both the benefactor and the recipient have the “correct” intentions given the situation in question. This subsection describes problems stemming from each of these conditions.

#### III.2.a Intentions of the benefactor

Consider Bertrand Russell's Inductivist Chicken (which soon became an Inductivist Turkey), as recounted by Alan Chalmers:

> The turkey found that, on his first morning at the turkey farm, that he was fed at 9 a.m. Being a good inductivist turkey he did not jump to conclusions. He waited until he collected a large number of observations that he was fed at 9 a.m. and made these observations under a wide range of circumstances … Each day he added another observation statement to his list. Finally he was satisfied that he had collected a number of observation statements to inductively infer that “I am always fed at 9 a.m.”. However on the morning of Christmas eve he was not fed but instead had his throat cut.\(^{105}\)

The parable is a critique of inductionism but it is also worth noting for our purposes that turkeys that are being raised for slaughter should not be grateful to farmers.\(^{106}\)

---

106 All else being equal, that is. Enterprising moral philosophers can surely find exceptions. For instance, I routinely ask my students to gauge the moral permissibility of someone adopting a dog that is just about to be put to sleep with the proviso that the adopter insists on killing the dog.
Less dramatically, we usually do not feel that someone owes a debt of gratitude toward a person who has benefited her inadvertently, indirectly, or as part of a self-serving course of action. The benefits must be conferred at least in part for the sake of the recipient, and the recipient's having benefited should be understood to be good in its own right. What's more, as Simmons points out, the effort made by the benefactor should count for something.¹⁰⁷

The distinction between psychological and moral gratitude is relevant here. I may feel grateful when someone's action fortuitously helps me or when at very little cost she saves me from dire trouble, but this does not mean that any duty has been created. Again, the feeling may be real; we may feel grateful or something very close to gratitude when we receive what for someone else is a pittance to give, even as we know that we have incurred no additional duties. As anyone who has witnessed or participated in a codependent relationship would know, sometimes we must remind ourselves that our intuitions are not always synchronized with our duties.

It is unlikely that the state provides benefits with the right intentions as far as a gratitude theory is concerned. Assuming that a just liberal state could exist, it would seem to have to be an institution whereby people collectively provide benefits to themselves.¹⁰⁸ This is, after all, close to the heart of self-rule. Whether the government follows a trustee or delegate model, appeals to gratitude seems odd.¹⁰⁹ For instance, if (painlessly) after a year's time. Assuming that dogs can have duties of gratitude and cannot feel dread, perhaps it would have a duty of gratitude toward the person who gave it an extra year to live. ¹⁰⁷ Simmons, Moral Principles, 170.
¹⁰⁸ For a more cynical vision of the emergence of the public provision of goods, see Olson, “Dictatorship, Democracy, and Development.”
¹⁰⁹ For a discussion on delegate vs. trustee models, see Pitkin, Concept of Representation. Delegate...
the state is essentially an amalgamation of delegates, then it is simply executing the orders of the people by using the resources of those people. How can we be grateful for executives simply doing their jobs and no more?

Even if the state apparatus looks more like a group of trustees, it is still (usually) providing services with the resources of its citizens.\textsuperscript{110} We may feel grateful if the state is especially effective, but this psychological gratitude is not the same as moral gratitude. To the extent that the state is a separate entity from the citizenry and is carrying out its assigned tasks, it is doing its duty. That is in effect the least it can do, morally speaking. We may be grateful that the state apparatus is composed of such competent people, but we would not be grateful towards the competent people because their efforts yield surprisingly positive results unless we found out that they worked exceptionally hard to produce those results.

\textbf{III.2.b Intentions of the recipient}

Perhaps the most important problem from the perspective of the recipient comes from Nozick's discussion and rejection of fair play theory, remarks that pertain here as well. Nozick argues convincingly that we cannot come to have new duties as a result of goods being foisted on us.\textsuperscript{111} The problem lies in the distinction between accepting and receiving goods, a distinction that I will discuss in more detail in the

\textsuperscript{110} Maybe the citizens of a state would have stronger duties of gratitude if the state sacked and pillaged foreign lands in order to spend more on its citizenry.

\textsuperscript{111} Nozick, \textit{Anarchy, State, and Utopia}, 90-95.
following chapter. Simmons argues that for a duty of gratitude to obtain, the recipient must first actually want the object or service given to her (although I should point out that she does not need to want it prospectively) and second, that she want it via her benefactor. 112 My teetotaling friend does not incur a debt of gratitude if I buy him a drink, especially if I know that he never consumes alcohol. I, on the other hand, might incur such a debt if my friend buys me a drink, but not necessarily if a stranger does so – especially if I doubt his or her intentions. Many regular bar-goers are rightly wary of strangers bearing gifts.

But the problem of intentionality on the part of the recipient runs deeper. As I mentioned earlier, the state (or its representatives) usually does not provide benefits for reasons that would engender duties of gratitude. The flip-side of this is that we also do not receive benefits in a way that would do so. Simmons has argued – plausibly, I think – that many citizens see themselves as consumers who are buying products with their tax dollars. 113 These “purchases” are not necessarily self-serving; those with money could happily support wealth transfers to the poor. Yet these are understood to be purchases nonetheless. So even if the state were somehow providing benefits for the “right” reasons, (1) citizens would likely not even know that they were now “in debt” 114, and (2) citizens could reasonably say that they never intended to accept benefits above and beyond what they paid for. 115 Again, in such a case they may have

112 Simmons, Moral Principles, 177-178.
113 Simmons, Obligations of Citizens, 57.
114 Simmons, Moral Principles, 181-182.
115 This is one of the theoretical downsides of abandoning the filial approach to gratitude. The distinction between accepting and receiving once again rears its head.
a duty to obey the law, but this duty would not be based on gratitude.\textsuperscript{116}

It is worth pointing out that transactional gratitude theory in particular does not have ready recourse to hypothetical consent. Recall that under filial gratitude theory, parents can plausibly say that their children do not yet know or understand why they – the parents – must provide some services even in the face of the children's objections. If we abandon this analogy, however, the state can no longer resort to paternalism to argue that (1) it must provide the citizenry with certain benefits in spite of its objections, and (2) that the citizenry, therefore, has a debt of gratitude toward the state which must be discharged by way of law abidance.

Finally, I should point out that it is possible for a philosophical anarchist to claim that social groups, or even society writ large, is owed a debt of gratitude. I do not share this position, but I also do not want to preclude it. But for this debt to become a debt to obey the state, more work needs to be done to show that the state is an appropriate embodiment of the relevant people or will of the people. What’s more, as I will discuss below, we need to know why we must repay “society” by obeying the law, and not by creatively trying to help others in idiosyncratic ways.

\textbf{III.3 The Value of Gifts}

We know that we should remove the price tag when we give a gift. Failing to do so sends the wrong message: that we want the recipient to be aware of exactly how much she owes us. Yet repayment in some form is required, and the conceptual

\textsuperscript{116} I will discuss collective goods more fully in the following chapter, where I discuss fair play theories.
problems surrounding just how much must be paid back further depends on the relationship between the giver and receiver.\textsuperscript{117} Although (or because) expressing gratitude can be usually be done in many different ways, and because what is considered appropriate depends on the relationship between the parties, demonstrating gratitude is far more complicated than fulfilling one's end of a bargain in an explicit contract.

The first difficulty is that without knowing what or how much we owe, it is not clear how we should repay our debt. For the sake of analytical simplicity, we might say that debts of gratitude occur as a result of specific actions. This seems to be in the background of the transactional understanding of duties of gratitude. Yet this simplification goes too far. Feelings of gratitude are frequently vague and diffuse.\textsuperscript{118} Those who would reply that duties, unlike feelings, are not vague, must confront two difficulties. First, when many people participate in the provision of the benefit, it is difficult to know what one owes to whom. I may owe my extended family a debt of gratitude, but it must be repaid toward individual members. And what (if anything) do I owe Cousin Joe who stops by once a year to make awkward small talk over Thanksgiving dinner? Second, it's not always clear what I owe. Social conventions help here – I may "know" that asking someone out for coffee in order to talk about my work means that I should probably offer to pay the bill, but these conventions are

\textsuperscript{117} Gift-giving presupposes reciprocity but not necessarily egalitarianism. For instance, Alan Fiske describes four types of relations – Communal Sharing, Authority Ranking, Equality Matching, and Market Pricing – each with different rules of reciprocity. I will discuss Fiske in greater detail in a later chapter. For a schematic overview of Fiske's model, see Fiske, \textit{Structures of Social Life}, 41-49.

\textsuperscript{118} Simmons, \textit{Moral Principles}, 168.
especially unlikely to emerge in the modern, pluralistic societies that so worry legitimacy theorists.

Attempts to codify these social conventions in order to have a stable way of determining duties either fails or moves us away from gratitude theory. One possibility is that we “just know” what's fair and that we therefore know what constitutes appropriate compensation for just or unjust actions done toward us. John Cacioppo and William Patrick have suggested as much in *Loneliness*, tying our ease at recognizing appropriate just desserts to deep biological urges for affiliation. This answer is problematic, however. First, as societies become more diverse, we (1) have more opportunities to make mistakes, as I will argue in my chapter on etiquette and (2) are called upon to repay debts of gratitude through in increasingly formal ways, ways that depart significantly from the implicit judgments that we make when assessing fairness.

Second, to the extent that the provision of benefits and reciprocal demonstrations of "gratitude" are specified, they fall out of the realm of "gratitude." At one extreme, we enter the realm of explicit contracts. Moving away from this and into the realm of implicit expectations, the recipient's duties look far more like duties of fairness, fair play, or respect than duties based on gratitude. The philosophical anarchist challenges to those theories of legitimacy would thus apply here as well.

This is a major reason why we may feel uncomfortable with the assertion that we should repay our debt of gratitude by obeying the law. This exhortation relies on

the misconception that benefactors can dictate the terms of repayment, as in “I did you a favor, so now you have to pay me back in the following way...” But as both Simmons and Gans point out, that is simply not how debts of gratitude work. Gratitude does not mean that the person giving the gift may dictate what is to be received in return.\textsuperscript{120} To the extent that the “correct” repayment has been socially determined, the debt is not one of gratitude. And when the repayment has not been pre-determined, debts of gratitude allow the grateful person leeway in how she demonstrates this gratitude.\textsuperscript{121} The philosophical anarchist could sensibly say that she should be able to express her appreciation for the state by furthering what she sees as its substantive interests without committing to obeying its laws. She could, in other words, discern the intent behind the laws – national security, privacy, aid to the poor – and work towards those aims without necessarily obeying the law.

I will try to illustrate this idea by way of an example from the field of computer security. Computer hackers are oftentimes divided into three groups.\textsuperscript{122} Two of these are straightforward. White hat hackers are employed by governments or private security firms and use their skills to safeguard information and data integrity. Black hat hackers use their skills to damage networks or defraud others.

\textsuperscript{120} Gans, \textit{Philosophical Anarchism}, 46; and Wellman and Simmons, \textit{Is There a Duty}, 119-120.

\textsuperscript{121} Of course, there are nonetheless limits to the types of suitable repayment, and this will be influenced by the type of relationship that exists between the participants. In the abstract, we can say that repayment with an identical object to the one received is usually offensive. Paying back too much or too little can also cause problems, notably embarrassment on either side. This type of embarrassment is morally significant in part (but not only) because it signals that one or both parties misunderstood the nature of their relationship, and the parties must now do work to re-establish some mutual understanding as to where they stand in relation to each other. I will have much more to say about this in subsequent chapters.

\textsuperscript{122} Zetter, “Hacker Lexicon.”
Grey hat hackers are those hackers who commit crimes without the destructive intent of black hats. They may, for instance, break into a computer network to better understand its (with no intention of using the information for financial gain, or for helping others gain or lose financially), or simply as a challenge. In fact, grey hats have been known to hack into systems and e-mail the system’s administrator to notify her of the breach, along with advice on how to fix it.

This latter type of activity is clearly illegal, but it is plausible for a philosophical anarchist hacker to reason that her illegal activities better serve the intent of privacy laws than would her strict law abidance. Were she to refrain from hacking, the security hole would likely remain and be exploited by people who would be happy to commit egregious privacy violations by taking the data and either making it public or selling it to the highest bidder. So even if the hacker were persuaded that she had a gratitude-based obligation toward the state, she could reason that given the relative technological illiteracy of legislators, some laws should be disregarded in order to fulfill that obligation.

One could argue that discharging our debts of gratitude toward the state are exceptional in that the type of repayment must consist of law abidance for the state to survive. To use an analogy: I can discharge my debts of gratitude to my well-off friends in any number of ways, but repaying a starving person by giving them free chess lessons misses the mark.\textsuperscript{123} Similarly, the argument goes, trying to advance some of the state's substantive interests, while possibly well-intentioned, would fail to

\textsuperscript{123} This is very similar to the position of natural duty theorists who argue that we have a duty to build and uphold just institutions. I discuss these in Chapter 5.
give the state what it actually needs.

Again, this reply is unpersuasive. Simmons correctly points out that even if the state requires obedience to survive, this does not imply that I must strive to obey all laws at all times. As others have argued, the state can absorb a certain amount of disobedience without falling apart or even being noticeably less effective. Therefore, even if I have a gratitude-based duty to obey, I may be able to discharge this duty by selectively obeying laws. In fact, if I have good reason to believe that law abidance in a particular case or class of cases would go unnoticed or not harm the state's substantive interests at all (that is, without interfering with the state’s ability to achieve what it’s trying to achieve by establishing laws), I might sensibly argue that discharging my duty by obeying the law would be tantamount to showing appreciation to a friend by buying her a record that she already owns. It would make more sense to try to discharge my duty in a way that would actually be helpful, even if this means breaking some rules. This, of course, is not what the archists have in mind. They would not be satisfied if philosophical anarchists conceded a duty to advance the state's substantive interests – or to at least avoid interfering with them – without the larger concession that they also have a prima facie duty to obey the law.

The problem of gratitude is complicated further by the fact that the continuous provision of benefits from the state means that we can never fully discharge our duties. In that sense, gratitude toward the state does seem closer to a filial than a transactional

124 Simmons, *Moral Principles and Political Obligations*, 186.
125 Wasserstrom, *The Obligation to Obey the Law*, 27-29.
126 And to be clear: philosophical anarchists would not make this concession.
relationship. Unless there are prior emotive connections that bind the citizenry, however, it would be a filial relationship devoid of the personal considerations we normally associate with these types of relationships. That said, positing the existence of these types of prior emotive connections would make other theories of political obligation less problematic as well, raising the question of why we should bother with a more fringe theory like gratitude in the first place.127

III.4 Goods relative to which bads? Gratitude and counterfactuals

Finally, gratitude theories usually do not compare the provision and reception / acceptance in one situation relative to counterfactuals. But doing so seems to me to be crucial when determining whether or not debts of gratitude exist at all. Even if we agree that duties of gratitude do exist in general and that a state unambiguously provides us with benefits, it does not follow that we owe that state our gratitude. For this, we would need a compelling account that the state provides us with sufficient benefits relative to a given baseline. There are four possible baselines: (1) some sort of universal minimal standard; (2) what a given state can plausibly provide; (3) idealized anarchist communities; and (4) anarchist communities embedded within a global state system. Each standard presents possible problems for a would-be gratitude theorist.

III.3.a Non-anarchist standards

127 To give one example, it would weaken the objections to utilitarianism in A Theory of Justice. For example, more prior emotive connections would mean more endorsement of utilitarian principles "for the right reasons."
First, consider the position that we owe the state obedience because it provides benefits relative to some sort of minimal standard. This view says that (1) our duties of gratitude are vitioted if the level of benefits falls below that standard and (2) we cannot avoid these duties if the benefits reach or exceed that standard. Yet it would be difficult to see how we could establish such a standard and why it would be compelling. If a state easily provides a minimal level of benefits but could provide far more, why should we be grateful for the proverbial crumbs off the banquet table, and would we not be entitled to wonder what is happening to the rest of the bread? Regimes that provide minimal benefits surely do not deserve debts of gratitude if they are otherwise plagued by corruption or incompetence.

What's more, we seem to have more of a reason to be grateful to those who help us when they do so at great cost to themselves. Once the benefactor reaches a certain level of affluence, our sense of gratitude becomes so diffuse that it is difficult to associate it with the benefactor as such. We may feel grateful for our luck of having been born into a position that allows us to benefit from the activities of the benefactor, but it becomes increasingly difficult to say that we owe duties of gratitude toward those who barely (or sometimes simply do not) realize that they are helping us. Simmons draws a similar distinction, differentiating between being happy with a state of affairs and being grateful to a benefactor. The mere fact that the state of affairs came about through the willful acts of others is not enough to generate duties of gratitude.

---
128 Simmons also makes this point. See Simmons, *Moral Principles*, 170.
129 Ibid., 172.
Perhaps then we should assess the worth of the benefits provided in terms of what we believe the state in question could provide. Poor states would not be faulted for meager benefit provision so long as these benefits exceed what would be available with (1) no state whatsoever, and (2) most other models of benefit provision available to that state. In this model, rich states or potentially rich states would not be excused for excessive waste or corruption. Yet even this promising strategy is subject to critique. While it may produce duties for many people, it would have difficulty creating general duties that would bind (nearly) all individuals. Duties of gratitude vary by individual, and the state's helping the poor does not inherently produce duties among the rich. In fact, many rich people would prefer lower taxes and lower transfers to the poor. The state is providing a level of services that is far below what the rich could expect under an alternative political arrangement, and as a result the rich would be excused from their political obligations according to this version of gratitude theory. Note that such transfers of wealth may or may not be justifiable; I am simply saying that it would be bizarre to say that the rich would have any duties of gratitude toward the state in this case.

This objection may be overcome on prudential grounds; the rich may sensibly believe that re-distribution of a portion of their wealth will lead to social peace or a thriving economy which ultimately benefits them. In such cases, they would indeed have a duty to obey based on gratitude. Yet unless everyone benefits more from the current state of affairs than they would under an alternative arrangement, many will have cause to complain that they do not owe a debt of gratitude because the state is not
benefiting them nearly as much as it could. In other words, this is an individuated version of the objection to the first standard. States constantly enact policies that hurt some at the expense of others, and even if such policies are defensible on the grounds of fairness or utility, we should not expect those who are severely harmed – relatively speaking – in the process to obey the law because they are grateful.

III.3.b Anarchist standards

But why should we restrict ourselves to comparing actual service provision with what states could provide? After all, even if the state is performing well relative to its potential, it may provide fewer benefits to many individuals than would be available in anarchic political arrangements. Ian Carter has observed that theories of legitimation such as Rawls' *Theory of Justice* take the state and an existing political community as given. But some may object that the existence of a political community as a stable, duties-producing entity is precisely what is in question. And if existing states or communities provide fewer benefits than anarchic arrangements, it once again becomes difficult to see why we would owe these states a debt of gratitude.

The final two possible standards that gratitude theorists might adopt take this objection seriously, although I will merely note the third and then discuss the fourth. The third standard is that the state provides significantly more benefits than we would have if no states existed. The fourth is that the state provides significantly more benefits than we would have if we lived in an anarchic area within a global state

---

130 Carter, “Evolution of Rawls’.”
system. The former takes us very far into a land of counterfactuals that draw deeply on imagination and the suspension of disbelief. Would a world without states make us think of Hobbes or Kropotkin? The prospect is so speculative that I have no way of knowing how to choose.

The fourth standard is more tractable, and archists might point to failed states like Somalia as evidence of the consequence of a lack of a robust state. This type of argument is disingenuous, however, when it relies on cherry-picking cases. First, anarchists have long argued that voluntary, fledgling anarchist communities have failed or been weakened due to deliberate pressure from established, state-based power structures. Would the Paris Commune have failed had it been left unmolested in its infancy? What about the nascent Spanish anarchist zones in the 1930s? Perhaps these communities would have collapsed or morphed into states; we just don't know. But if they were preferable to communities under states (all else equal) and they did fail as a result of interference by hostile states, then the solution is not to adopt an archist gratitude theory perspective, but rather to become a global political anarchist.

Archists might insist that the central purpose of the state is to protect its citizens from other hostile states – a variation of the NRA slogan that “the only thing that stops a bad guy with a gun is good guy with a gun”, and that in the absence of global anarchism, we should at least for now be grateful toward the state. Anarchists

131 For the sake of brevity, I will ignore the very plausible objection that we should compare current service provision with that which would take place if we were organized in ways that were neither anarchic nor state-centered. States, after all, are only one type of ostensibly authoritative political structure.

132 Overby, “Bad Guy With a Gun.”
could reply that first, there is no a priori reason to suspect that self-defense can only be undertaken via a centralized, authoritarian institutional structure and second, that it would perhaps be best if other states did take over one's own. In other words, given the reality that they will live in one state or another for the foreseeable future, anarchists might feel that being "conquered" by a less awful state is the lesser of two evils. In such a case, gratitude theory would mean defending one's state even if doing so was against one's interests and the interests of one's community or citizenry as a whole.

What's more, some promising anarchic communities exist in spite of the hostility of states to their existence. Raúl Zibechi has chronicled a number of communities in Latin America which have responded to neoliberal restructuring by developing self-regulating communal zones. As unions collapsed and no longer seen as effective vehicles to petition the state for additional resources, leftists switched strategies. Instead of engaging the state, they increasingly either ignored it – treating it as irrelevant – or, when confronted by government officials, actively resisting it. These are not small fiefdoms; they are highly participatory communities that stand in contrast with the hierarchical provision of law, order, and other serves that we associate with the liberal state.

There is, unfortunately, a dearth of reliable data on the outcomes of these recent experiments, but by at least some accounts these people seem to prefer these self-regulating communities to living their lives under the state. If those living in these ostensibly autonomous communities (1) prefer doing so to living under a normal

133 Zibechi, Territories in Resistance.
government, and (2) are not doing so in violation to some prior principle of fairness, I do not see why those members would owe the types of debts of gratitude to the state that would satisfy archists.

One might complain that if we grant that these smaller, purportedly anarchist communities could exist and could provide goods far more effectively than nation-states, then we have merely moved the locus of legitimacy down to the local level. In other words: if an autonomous zone in Peru provides more services than does the incompetent or uncaring federal government, couldn’t we just say that the inhabitants of that zone have a duty of gratitude towards whatever organizational apparatus runs that zone, or to that zone as a whole?

The difference between the state and an anarchist community may not be as stark as meets the eye. Weber describes the state as “a human community that (successfully) lays claim to the monopoly of the legitimate physical violence - and this idea of ‘territory’ is an essential defining feature.”\(^{134}\) This definition clearly applies to modern states as we understand them, but there is no prima facie reason why this description could not apply to anarchist zones. These zones do use force to mete out justice and apparently reject any outsider’s right to do so.

Yet leaving aside local communities that merely recreate basic state structures (as we colloquially understand them) on a more local level, applying the definition to autonomous zones is problematic in at least two ways.

First, the nature of anarchist zones eliminates the distance between government

\(^{134}\) Weber, “Politics as a Vocation”, 33.
and citizenry that forms the basis for gratitude theory. Anarchist societies are participatory; if they weren’t, they wouldn’t be anarchist. In those cases, the idea that I would be grateful to an entity that provides services is undermined by the fact that, unlike in representative democracies, I am a direct part of that entity. This lack of separation makes a fair play or natural duty theory of obligation far more intuitive than a gratitude theory when it comes to anarchist communities.

Second, the Weberian definition of the state is undermined in anarchist communities insofar as the communities (1) do not claim a long-term monopoly of power over a stable piece of territory, and (2) accept that, as per Godwin (among others), anarchist communities should be comprised of a federation of organizations where no group can claim sovereignty and where the organizations (a) can associate, disassociate with, and merge and separate from each other relatively easily, (b) can change jurisdictional or territorial boundaries as per the desires of their membership, and (c) allows for an easy inflow and outflow of members across groups. In such a truly anarchic context, to which frequently morphing groups should one’s gratitude be directed? By their participatory nature and fluidity, political anarchism makes the appropriate recipient of gratitude a moving, morphing target. Gratitude theory appears problematic regardless of whether it is applied to the state or to an anarchist zone.
Chapter 3: Fair Play

I. Introduction

Having discussed theories of obligation based on gratitude, I now turn to fair play theory. After describing its evolution, I concentrate on what the argument George Klosko makes in *The Principle of Fairness and Political Obligation*. This, I believe, is the strongest articulation of fair play theory, but I ultimately find it insufficient. More specifically, the theory is insufficient in itself and must be buttressed by another theory of obligation. The need for this support raises the question of whether or not we can simply dismiss fair play theory altogether and adopt the “supportive” scheme as a general theory of obligation.

The remainder of this chapter is divided into five sections. Section II first explains early statements of fair play theory propounded by Hart and (pre-Theory of Justice) Rawls and then presents existing arguments against these incarnations of the theory. With these deficiencies highlighted, Section III explains Klosko's "reloaded" version fair play theory, one that tries to keep the good in Hart and Rawls while shoring up the weaknesses of their accounts. Section IV is a critique of Klosko’s scheme. In section V, I discuss an alternative argument against fair play theory based on Jeffrey Green’s recent defense of “plebian democracy.” Section VI is a very brief conclusion.

II. The Hart-Rawls Theory and Its Challenges

II.1 Original formulation
If we tried thinking up a slogan to attach to fair play theory, one strong contender would surely be: “Don’t mooch.” H.L.A. Hart’s 1955 piece *Are There any Natural Rights?* is widely seen as the seminal formulation of modern fair play theory. His definition, more elaborate than my slogan, is as follows:

When a number of persons conduct any joint enterprise according to rules and thus restrict their liberty, those who have submitted to these restrictions when required have a right to a similar submission from those who have benefited from their submission. The rules may provide that officials should have authority to enforce obedience… but the moral obligation to obey the rules in such circumstances is due to the cooperating members of the society, and they have the correlative moral right to obedience.\(^{135}\)

Following Hart, a young John Rawls followed suit and proposed a very similar sounding theory in his 1964 paper, *Legal Obligation and the Duty of Fair Play*. In it, he says that

Suppose there is a mutually beneficial and just scheme of social cooperation, and that the advantages it yields can only be obtained if everyone, or nearly everyone, cooperates. Suppose further that cooperation requires a certain sacrifice from each person, or at least involves a certain restriction of his liberty. Suppose finally that the benefits produced by cooperation are, up to a certain point, free: that is, the scheme of cooperation is unstable in the sense that if any one person knows that all (or nearly all) of the others will continue to do their part, he will still be able to share a gain from the scheme even if he does not do his part. Under these conditions a person who has accepted the benefits of the scheme is bound by a duty of fair play to do his part and not take advantage of the free benefits by not cooperating.\(^{136}\)

The first thing to notice about these accounts is that they are very, very similar. In fact,

---

they are generally acknowledged to be functionally interchangeable, with one exception which I will discuss later. Simmons argues that there are three key aspects common to both theories, and to the best of my knowledge his characterization has never been seriously challenged. First, in both schemes there must be an active scheme of social cooperation. These schemes must be mutually beneficial, a point made explicit in Rawls and taken to be implicit in Hart. Second, cooperation under the scheme involves a restriction of one’s liberty. Third, the benefits yielded by the scheme may be received in some cases by someone who does not cooperate when his turn comes. Given these similarities, I, following others, will lump together these two formulations and refer to them collectively as the “Hart-Rawls scheme.” The salient difference between Hart and Rawls that I alluded to earlier is that Rawls adds the further stipulation that the scheme must be "just." While I treat the schemes as interchangeable for the moment, this difference will become relevant in section IV.1 of this chapter, where I present my first objection to Klosko’s scheme.

Before continuing, I want to flag a couple of crucial features of the Hart-Rawls scheme. First, notice that it is “closed”; the obligations are generated because a finite number of people are involved in a joint project, and a certain finite number of people can take advantage of the fruits of the labor of those who are restricting their liberty. Whatever obligations I have toward those who are not involved in the scheme are obligations that derive from something other than fair play theory. I ostensibly have duties of fair play to fellow Americans, but my duties to world's poor may be based on

137 Simmons, *Moral Principles*, 104-106.
a duty of charity, respect for human rights, and so on. Second, it is important to keep in mind that fair play theories are a distinct subset of fairness theories. Many competing theories make claims about fairness. There is nothing fair about my being born into a society where I had easy access to higher education, did not face discrimination based on my race or gender, had enough resources to have the luxury of studying political theory, etc., when so much of the world’s population does not share these advantages. This is true, but it also not the kind of fairness discussed in Hart-Rawls. To put it simply: if there is no joint project, it is not a fair play theory of obligation.

II.2 Objections to the original formulations

Enthusiasm for the fair play theory of obligation did not last long. Perhaps seeing the writing on the wall, Rawls quickly turned his back on it. The problem, as he explains later in Theory of Justice, is that it seems wrong to suggest that actual political societies are made up of people who freely accept the benefits of cooperative schemes.\textsuperscript{138} As Rawls puts it:

\begin{quote}
\textit{...citizens would not be bound to even a just constitution unless they have accepted and intend to continue to accept its benefits. Moreover, this acceptance must be in some appropriate sense voluntary. But what is this sense? It is difficult to find a plausible account in the case of the political system into which we are born and begin our lives.}\textsuperscript{139}
\end{quote}

With this in mind, Rawls transitions to a duty to uphold just institutions. “Theory of

\begin{footnotes}
\item[139] Ibid., 336-337.
\end{footnotes}
*Justice* Rawls,* then, is grounding political obligation on "justice as fairness," but this fairness is not the same fairness as the one that constitutes fair play theory.

Following Rawls, Nozick’s *Anarchy, State, and Utopia* provides a series of famous hypothetical examples designed to highlight deficiencies in (Hart’s) fair play theory. The most cited example can be summarized as follows. You are a member of a community of 365 adults. Some members - call them “doers” - decide to set up a public address system in order to provide entertainment for the citizens. Since each citizen benefits from this arrangement, the doers decide that every member of the community should spend a day providing content for the other members. They draw up a list of everyone’s name and assign each person a date. What happens when your date comes? Nozick puts it this way:

> Are you obliged to take your turn? You have benefited from it, occasionally opening your window to listen, enjoying some music or chuckling at someone’s funny story. The other people have put themselves out. But must you answer the call when it is your turn to do so? As it stands, surely not. Though you benefit from the arrangement, you may know all along that 364 days of entertainment supplied by others will not be worth your giving up one day.140

The thrust of the argument is simple but profound: we can’t be roped into a fair play scheme against our will.

Yet Nozick's critique has itself been subject to various critiques. Providing a catalog of objections would lead me too far astray, but A.J. Simmons' critiques are salient here as they sharpen our focus on where exactly the problem lies. While

---

Simmons agrees that the Hart-Rawls scheme does not work, he is nonetheless more charitable than Nozick. Simmons’s central insight here is based on two distinctions that he makes. The first distinction is between accepting and receiving benefits or goods. For Simmons, to accept a benefit it to either try (and succeed) in getting the benefit or to benefit from something or some action willingly and knowingly. If a good is not accepted but I still “have” it, then it is merely received. The second distinction is between open benefits and readily accessible benefits. Open benefits are those that I can’t avoid receiving “… without considerably altering my lifestyle. Otherwise, benefits are readily accessible.”

These distinctions should help make Simmons’s problem with Nozick clear. Simmons believes that to accept a good quite clearly generates obligations of fair play. Simply put, if either condition for acceptance holds in my case, then I’m mooching. “Readily accessible” benefits are the types of benefits that must be accepted if one is to have them at all. Therefore, in the case of readily accessible benefits, objections to the Hart-Rawls scheme fail. Fair play duties do obtain in such cases, and Nozick is wrong to make general statements against fair play theories. The real problem with fair play lies with cases involving "open benefits." If I am merely receiving these benefits, then fair play does not generate duties to "do my part." I must accept them for those duties to kick in. This is no trivial matter; benefits like “security” count as open benefits.

141 Simmons, *Moral Principles*, 129.
142 Ibid., 130.
Crucially, Simmons argues that most people do not accept these benefits.\textsuperscript{143} According to him, many of us barely notice the benefits we receive. To the extent that we do see benefits, we do not see them as the result of a cooperative venture; we see them as having been purchased like any other commodity or service. What’s more, he argues that many citizens simply do not believe that the benefits are worth the price. Like Nozick’s (non) free-rider who enjoys the public address system but doesn’t think it’s worth the cost, many of Simmons’s citizens like the open benefits but don’t think that they are a good value.

If Simmons is right, then fair play duties obtain if and only if a person has accepted (and not merely received) benefits. This is especially a problem in the case of open benefits. Simmons’s account thus turns heavily on the psychological traits of those who are benefiting.\textsuperscript{144} What a beneficiary thinks is happening will at times make a decisive difference. Turning to Klosko, then, we can see that his central task will have to be to provide some way of grounding a duty of fair play in the case of open benefits that have not clearly been "accepted."

\section*{III. Klosko’s Theory}

\subsection*{III.1 Klosko's Three Criteria}

As should be clear from the preceding section, the main task of any fair play

\begin{flushright}
\textsuperscript{143} For an especially good discussion of this, see Menlowe, “Political Obligation.” Simmons also discusses this in \textit{Moral Principles}, 138-139. \\
\textsuperscript{144} I should point out that Simmons’s “psychological criteria” (e.g.: understanding benefits as benefits, being aware of benefits) is itself up for debate. See especially Richard Arneson's sustained critique, in Arneson, “Principle of Fairness.”
\end{flushright}
theory must be to explain how obligations are generated in the case of a scheme providing what are “open” goods. So long as these cannot be readily accounted for, fair play theory will not have much bite in terms of justifying a general obligation to obey the law. Klosko takes up the gauntlet, most notably in his *The Principle of Fairness and Political Obligation* (1992), and argues that the principle of fairness can in fact apply in cases of open benefits if each of the following three conditions are met: (i) the goods must be worth the recipients' effort in providing them; (ii) the goods must be "presumptively beneficial"; (iii) the benefits and burdens must be fairly distributed.145

All of these conditions are necessary, but it is the "presumptively beneficial" criterion that does much of the heavy lifting - it will turn out that this is what will ostensibly justify the provision of many goods that are not at all presumptively beneficial. What makes a good a "presumptive good"? Klosko acknowledges that what constitutes a presumptive good is "somewhat vague and difficult," but it turns out that presumptive goods are public goods that "... are necessary for an acceptable life for all members of the community."146 More precisely, they are goods that we believe are necessary, although this belief can be revised.

Ignoring for the moment the work that presumptive goods will later do in his theory, we can immediately see that a good's being presumptive will go a long way toward satisfying Klosko's first criterion - that the good must be worth the recipient's

---

146 Ibid.
effort. In fact, if something is a presumptive good with respect to me, then the first criterion seemingly must be satisfied in my case. After all, if the good is necessary, then surely it is worth the effort unless this effort involves losing or not obtaining another presumptive good.

What is not clear in Klosko’s account is whether we should think of presumptive goods as pertaining to specific individuals or people in the abstract. What I mean is that two interpretations seem consistent with Klosko’s use of the term: (1) that a good is presumptive to me if it is, in fact, necessary for me to live an acceptable life, or (2) what we call “presumptive good” is the type of good that we have good reason to assume is something that a typical person in a community needs in order to live an acceptable life. The latter formulation seems more palatable as it seems more compatible with Klosko’s claims. Klosko rightly points out that it is possible that some goods that we think of as presumptive may not be presumptive in some cases for particular individuals.147

This is what differentiates Klosko from a hypothetical consent theorist, i.e., someone who would say that presumptive goods are to which one should consent in order to further one’s interest..148 To appreciate the difference between Klosko and a such a theorist, consider the divergent reactions each theory would have to my saying “No, I don’t actually want that good, at least not for that price.” A proponent of hypothetical consent would reply with something like "You are mistaken in spite of

---

147 Ibid., 49, 116.
148 Again, see Pitkin for more on hypothetical consent.
yourself," or "You cannot consistently say that," or "You cannot qua rational human being say that." A proponent of Klosko's scheme, on the other hand, would likely raise her eyebrows and say "Are you sure? That sounds very far-fetched. Have you really thought this through?" In other words, Klosko allows that these particular individuals could in principle demonstrate or at least argue that the presumptive good in question is not necessary for them to live an acceptable life after all. The burden would be on these odd characters to provide such an explanation, but an acceptable explanation is possible. That said, these exceptions do not seriously undermine Klosko's framework if we take Simmons’ three criteria at face value because Klosko believes there just aren't many of these "weird" individuals. So long as Klosko's scheme applies to a great majority of individuals, it would not run afoul Simmons’ criterion that an acceptable theory of legitimacy “capture” the vast majority of those it should capture.\(^\text{149}\)

Klosko's third criterion, that goods be distributed fairly, maps onto the intuition that some people should not have to contribute disproportionately to a given scheme. Even in the case presumptive goods, I should not have a duty of “fair play” if I’m required to do all of the work while other beneficiaries play ping-pong. The problem, as Klosko recognizes, is that "fairness" is a contested concept – unless we insist on equal contributions, how do we determine when the required contribution becomes “disproportional”? Does the billionaire with a slightly higher marginal tax rate have a case when she complains that she is asked to make “disproportional” contributions?

\(^{149}\) It would however fail to resolve the problem that I introduced in Chapter 1, that Simmons doesn’t provide a reason why a non-consentor should care if Simmons’ three criteria are satisfied.
Klosko’s strategy here becomes puzzling. First, he decides to reject the goal of finding a fair distribution of goods and burdens and instead advocates using a fair procedure to do so. So far so good. To the obvious charge that there will be disagreement about what constitutes a fair procedure, Klosko replies that:

We should be able to select an acceptable (emphasis his) principle of fair distribution through procedures that are reasonably fair, even if they are not perfect. Though the notion of procedural fairness is essentially contested, proponents of different particular views would likely agree on certain aspects of decision mechanisms. Perhaps the clearest aspect is that an acceptable procedure should be democratic, granting each individual the right to have his or her opinion considered.¹⁵⁰

This is itself unsatisfying, and I will take up some questions that it raises later.

So far, we have a proceduralist account. But Klosko does not want a pure proceduralist account; shortly after dismissing the effort to find a fair distribution of goods, he says that the outcomes of the fair procedure in question must "be defensible with reasonable arguments."¹⁵¹ The most plausible way I can interpret this is that we should understand him to say that both the procedure and its outcomes must be "fair." Perhaps Klosko implicitly wants to privilege procedure over distribution by putting procedure earlier in the sequence of relevant considerations. Putting words in his mouth, maybe he wants to say: “If the procedure is unfair, stop and select a fair procedure. If the procedure is fair, then observe the distribution and ensure that it is defensible with reasonable arguments.” Since he does not seem to be explicit on the

¹⁵⁰  Klosko, Principle of Fairness, 65.
¹⁵¹  Klosko, Principle of Fairness, 66.
matter, I am unclear as to what to do in situations where the distribution is judged to be unfair. Do we repeat the same procedure and hope for a different result, or do we select a different “fair procedure”?  

Here another caveat appears: what counts as fair is to be understood in terms of an acceptable level of unfairness. As Klosko puts it: "As long as the unfairness of procedure X is not excessive - as long as it does not exceed the 'limits of tolerable injustice,' to use Rawls's phrase - the procedure can still be acceptable, though it is not completely fair." Given his earlier comments, we can assume that this standard also applies when we evaluate the outputs of this "procedure X." At this point, however, Klosko asks us to accept that we won't be able to draw a clear line between fair and unfair procedures. This does not seem to bother him, however, as he has strong faith - a faith that he seems to assume that his readers implicitly share - in the legitimacy of modern liberal democracies: “Strong evidence of the fact that the procedural mechanisms in existing liberal societies are generally viewed as acceptably fair is the fact that the means through which controversial issues are dealt with, and their outcomes are generally accepted." If I were to complain that I would prefer another method of providing a given presumptive good, Klosko would reply that this preference is insufficient to excuse me of my obligations. So long as the winning method has been chosen via a reasonably fair method and conforms to criteria (i)-(iii),

---

152 It's worth pausing at this point to note that Klosko is here sliding into a different use of the word "fair." We need one type of "fairness" in order to understand and explain a different kind of "fairness," i.e., that found in fair play theory.
153 Klosko, Principle of Fairness, 66.
154 Ibid., 68.
I have an obligation to obey, according to Klosko.

### III.2 Discretionary Goods

If Klosko has been successful thus far, he has adequately argued that we have political obligations pertaining to the provision of presumptive goods. Yet a government restricted to providing only presumptive goods would seem very different from any American administration in recent memory. Libertarians would cheer this scheme, as it seems to justify a nightwatchman state and little else; security is clearly a presumptive good, but (for instance) public funding for education is not. Klosko thus has two choices. He can bite the bullet and go through what Nozick describes experiencing while writing Anarchy, State, and Utopia: “With reluctance, I found myself becoming convinced of (as they are often now called) libertarian views, due to various considerations and arguments.”\(^{155}\) Or he must find some way of justifying the collective provision of discretionary goods. Klosko is no libertarian and takes the latter approach. In doing so, he uses two general strategies. These he calls the *indirect argument* and the *institutional argument*. The indirect argument bears a disproportionate amount of the justificatory burden, so I will pay especially close attention to it.

Klosko characterizes the *indirect argument* in various ways, but the crux is as follows: "If cooperative scheme X provides both presumptive and discretionary public

---

\(^{155}\) Nozick, *Anarchy, State, and Utopia*, ix.
goods, then under certain conditions these goods comprise an indivisible benefit package.”\textsuperscript{156} And later, "... the crucial point is that presumptive public goods cannot, practically speaking, be furnished by some government unless a certain range of discretionary public goods is furnished as well."\textsuperscript{157} Simply put, you can't have an army if you don't have roads. You can't have a police and firefighting force if you don't have a supply of electricity, and you can't have that if you don't have a citizenry that has at least a minimum level of education, health, and so on. Therefore, Klosko argues, libertarians and philosophical anarchists like Simmons can't declare that they'll only contribute taxes that go towards the military and other institutions that directly provide presumptive goods, as doing so would unfairly underplay the importance of the infrastructure that is required to keep those key institutions running. The idea is that this supportive infrastructure can get very extensive indeed.

At this point, Klosko is subject to two rejoinders. The first is that the discretionary goods that are needed to produce presumptive goods could be produced privately. Higher education is a necessary discretionary good, but why can't the free market provide and maintain the necessary institutions to supply it? Perhaps the state would have to take over if it was shown that this sphere of economic activity was subject to repeated market failures, but this doesn't seem plausible in most cases. Surely the market could produce at least the minimum level of discretionary goods needed to support the provision of presumptive public goods.

\textsuperscript{156} Klosko, Principle of Fairness, 82.
\textsuperscript{157} Ibid., 89.
Klosko's reply here is that while this pro-market stance should not be rejected out of hand, it should be thought of as only one option out of a set of acceptable options. Speaking about a libertarian "L," he says that

L is surely entitled to his opinion, and if community X has a democratic decision process, he is also entitled to have his opinion considered along with other competing opinions. But because the X-ites are likely to disagree sharply about the package of discretionary goods that is required for the performance of X's presumptive functions, this questions should be settled through the employment of tolerably fair procedures.158

In other words, "libertarianism" is a distributional scheme that can be "defended with reasonable arguments," but so long as the antecedent decision procedure is fair, the libertarian in a welfare state cannot complain of injustice; she has simply been outvoted.

The second rejoinder is that there is a subset of discretionary goods that consists of the kind of things that many people (and apparently Klosko) want the government to provide, but that can't seriously be considered necessary in the provision of presumptive goods. One example would be funding for the arts. The idea that arts funding is somehow indirectly necessary for the provision of any presumptive goods lacks face validity. This is not to say that funding for the arts is frivolous. Perhaps it has real worth, both intrinsic and instrumental. What it does not obviously have is a clear connection to public goods that are presumed to be necessary for an acceptable life for all members of the community.

158 Ibid., 90.
To this objection Klosko makes a retreat while trying to mitigate the damage. His retreat consists in acknowledging that fair play theory as he presents it probably can't account for this subset of discretionary goods. In such cases, he believes that other theories of obligation - utilitarianism, natural duty, charity, etc. - can pick up the slack. I shouldn't feel obliged to support funding for the opera in order to avoid mooching; I should support it because (for instance) I have a duty to help those around me develop their talents and aesthetic sensibilities. Yet Klosko also reassures us that there may be fewer such cases than we realize. That is because, Klosko argues, institutions that are legitimately in place to provide either presumptive goods or clearly necessary discretionary goods could also be used to provide additional discretionary goods. We may not be entitled (qua fair play advocates) to construct the type of public address speaker system that Nozick discusses with the goal of providing entertainment, but perhaps such a system could be justified in terms of its ability to disseminate vital information in times of crisis. Once this emergency broadcast network is up and running, the theory goes, it could then be used to play (taxpayer funded) smooth jazz, so long as the decision to do so was made by way of a reasonably fair decision procedure.

So much for the indirect argument. Klosko’s *institutional argument* amounts to the following: "We can refer to the benefits received from law and order and the existence of a working social decision process as *institutional goods*. Because the benefits that A receives from institutional goods depend on widespread willingness to comply, according to the principle of fairness, he too has an obligation to
Surprisingly, Klosko cashes this out in what looks like distinctly utilitarian terms. First, according to this argument, "The main reason why individuals must obey the law is the corrosive effects of disobedience." Given that any disobedience "corrodes" respect for and adherence to the law, disobeying laws that do not contribute to the provision of presumptive public goods - either directly or indirectly - ostensibly weakens our commitment to laws that do contribute to presumptive goods. Disobeying laws that support art funding will make it more likely that I develop a generalized disposition against obedience, a disposition that may eventually lead me to disobey in cases involving matters of national security.

Klosko is effectively claiming that obedience in a particular case should be understood in terms of developing habits. One cigarette is unlikely to harm me, but one cigarette might lead to another. Just as this "harmless" cigarette may lead to disaster in terms of my physical health, this instance of innocent jaywalking may be perilous to my moral health. It's worth remembering at this point that Klosko is comfortable with the prospect that I may justifiably break the law. What he is claiming here is that the burden of proof lies on me whenever I decide to break the law. I'm the one who has to argue that my countervailing moral reasons are sufficient to overcome my obligation to obey the law.\(^{161}\)

The upshot of Klosko's link between presumptive and discretionary goods is

---

\(^{160}\) Ibid.
\(^{161}\) It's also worth pointing that while Klosko believes that all laws carry moral weight, they do not carry the same moral weight. Thus it may be that very little is required to overcome the weight of a bizarre, archaic law.
that when faced with someone who dissents with respect to a specific law, the majority can reply that they only need to justify a set of policies, not each specific policy within the set. This set of policies is one of a great many possible sets, and it is justifiable so long as it meets Klosko's three conditions. The fact that I would prefer another set does count for something inasmuch as I should be able to use fair procedures to try to enact my preferred set, but it is not weighty enough, Klosko argues, to overcome my duties to support fair practices that conform to the three conditions. True, the actual distribution of goods will never be entirely fair - hence the different health, educational, and social outcomes among different social classes - but Klosko believes that societies like the contemporary US satisfy a roughly "imperfect" but still acceptable fair distribution of goods.

In fairness to Klosko, what counts as "unfair" is a problem that is not unique to fair play theory. A proponent of fair play theory could rightly say that critics who latch on to the indeterminacy of "fair" should get their own house in order. Indeterminacy is present in any moral theory, and the critics themselves must face the following dilemma: set the bar for "fairness" too low, and you permit injustice. Set it too high and you set an impossible to reach, and thus irrelevant, standard. Klosko, far from ducking this dilemma, accepts it as a fact of life and tries to organize a system of justification while keeping the dilemma in mind.

IV. Challenges to Klosko

I have four critiques - or in some cases pointed questions - of Klosko's scheme.
The first is that the scheme can generate obligations in the case of evil but internally fair regimes. The second is that its reliance on other theories of obligation makes fair play redundant. The third is that the empirical data he provides is subject to different interpretations. The fourth is that the scheme may constrain us to offer support to regimes that satisfy Klosko’s checklist but still somehow seem quite unfair.

IV.1 Evil but fair regimes

The fact that a government supplies us with presumptive public goods does not prima facie entitle it to obedience. An evil regime does not become worthy of respect because it provides its citizens with such goods, even if it does so fairly and if it is producing outcomes that, compared to alternatives, are not especially bad to those who participate in the scheme.

Consider the following scenario. Joe is sent to jail for some impersonal money-related crime such as bank fraud. He has pleaded guilty and in his heart knows that he holds no ill will to any particular individuals or groups; he just really wanted the money and expected the bank to be able to absorb the cost. Once in jail, Joe, who is white, is approached by a white supremacist prison gang.\footnote{162 For what it’s worth, I got this idea while watching two prison documentaries, \textit{Inside the Aryan Brotherhood} and \textit{The Aryan Brotherhood of Texas.}} This gang tries to recruit him and asks that he render services to it. Joe is aghast and resists, whereupon one of the gang members informs Joe of some pertinent facts of prison life. The prison, it turns out, is effectively run by the various gangs, which are divided by race. The
harried and underfunded prison guards understand this and tacitly allow the gangs much leeway. These gangs occasionally skirmish but exist more or less in a rough state of peaceful equilibrium. That said, given the opportunity, one gang would quickly and ruthlessly annihilate the other. It is understood, the gang member continues, that new prisoners are to affiliate with a gang that matches their race. To not do so is to forgo protection and to court grave danger.

But maybe Joe is a courageous fraudster. He replies that he is no racist and that although he very much wants to live, he is willing to take his chances by going it alone. The gang member, undeterred, says that Joe is already benefiting from protection. Everyone knows that gangs communicate by observing each other's actions and treating these actions as signals. A gang’s failure to defend a vulnerable member of its race would be a dangerous bit of non-action that might signal the gang’s weakness and embolden the opposing gangs to try more attacks. If the white supremacist gang does not adopt a blanket policy of defense of white inmates, they risk annihilation and make a mutually destructive gang war much more likely. All gangs know that that attacking a member of another race is a very risky game and so generally do not do so. With the current system in place, there is no need to test the other gangs' resolve – it is simply assumed that retaliation will occur, so all gangs desist from instigating violence. Thus, Joe is being defended by the white supremacist gang's well-advertised policy of retaliation. He and all other prisoners benefit from the peaceful (if tension-filled and racially hostile) climate that pervades the jail.

In this case, it is hard to see how fair play theory would not apply to Joe. The
gang is providing a presumptive public good, i.e., physical security. Without it, Joe would be risking his life. Since it is providing a presumptive good, this would mean that the gang would be entitled to insist that he support other "goods" that they provide, whether or not he agree with what these goods actually are. Perhaps this involves helping the gang members with their paperwork, informally negotiating on their behalf with the guards, distributing racist paraphernalia, or helping provide a tacit threat when the gang extorts non-members. As a gang member, he would be able to voice his concerns but would be overruled by the vast majority of the gang's activist members, who then honestly declare that they believe the decision procedure (a vote) to have been fair.

Something seems wrong here. Given known facts about human frailty, it would be understandable if Joe helped the gang, but it's odd to say that he has a full-fledged duty to support, even indirectly, white supremacy. A proponent of fair play theory might reply that the specific acts that are being asked of him are substantively below some moral standard, and this makes the requests invalid. This reply would be insufficient. The reply holds that either (a) the specific actions Joe is being asked to undertake are evil or (b) Joe is contributing to an overall evil regime. Let us state by hypothesis that (a) is not the case. Perhaps the gang excuses him from acts that are themselves vile so long as he performs other acts that are not inherently evil but that are nonetheless supportive, such as helping gang members negotiate the appeals process, and which free up the gang members’ time so that they can undertake evil actions.
Yet (b) also seems problematic, and the problem is instructive. In discussing it, we should think back to the “justice criterion” (described in section II.1) that distinguished Rawls’s scheme from Hart’s. A defender of Klosko might appeal to the criterion to say that Rawls, but not Hart, has a scheme that generates duties based on fair play only if the regime is not unjust towards others. I will be following Simmons’s discussion of this criterion up to a point but will omit his objections to the criterion that are irrelevant to this chapter.

The fundamental problem with the justice criterion is that it does not seem logically constitutive of the fair play principle; it looks tacked on. As Simmons puts it,

> The special rights and obligations which arise under the principle are thought to do so because of the special relationships which exist between the cooperating participants; a fair share of the burdens is thought to be owed by a benefiting participant simply because others have sacrificed to allow him to benefit within a cooperative scheme. No reference is made here to the morally acceptable status of the scheme.\(^{163}\)

In other words, something’s being unjust or evil may give me a reason to not perform an action, but it is conceptually separate from the logic of obligation that arises from fair play. It amounts to saying that I do have a duty to the white supremacists and to uphold their racist projects (even if these projects are not directly tied to the gang's providing me with security) but that this duty is outweighed by other considerations. That still seems odd. Contrary to this interpretation, Joe shouldn’t be faulted for saying that he feels no dilemma in disobeying the gang.

---

Simmons offers another interpretation of the criterion. He says:

We suggest, first, that what the justice condition does is, in effect, amend the principle to read that a person is bound to do his fair share in supporting a cooperative scheme only if he has been allocated a fair share of the benefits of the scheme. [...] In fact, this may be the argument that Rawls is suggesting when, in elaborating on the principle, he notes that if the scheme is just, ‘each person receives a fair share when all (himself included) do their part.’\footnote{164}

Simmons has a number of objections to this, but my point here is that this interpretation points to a Rawls who, like the Rawls in \textit{Theory of Justice}, abstracts away from the relationship between the group and other groups. In other words, it is possible to have a group that provides presumptive goods for its members and that uses internally fair decision-making procedures while at the same time doing horrible things to non-members. It is not at all clear how fair play proponents can respond to this other than by a fiat that says: “\textit{If your scheme does bad things to other groups, then your obligation is void.}” But this is unsatisfying; why should it be void if the goods are provided and the procedures fair? If Joe doesn’t like this, he can, like Klosko’s libertarian, voice his concerns and vote. So long as the gang allows him to do so\footnote{165} and can provide reasonable justifications for the distribution of their goods (and presumably their policy toward outside groups?), Klosko should have to say that Joe has a duty to obey.

\textbf{IV.2 The reliance on other theories}

\footnotetext{164}{Ibid., 111-12.}
\footnotetext{165}{In fact, many (often racist) motorcycle gangs do in fact do so.}
While some would argue that this is not a problem per se, Klosko may find it difficult to escape the charge that his scheme does work, but only insofar as it supports a minimalist state or something very close to it. As I mentioned in the section on the provision of discretionary goods, Klosko is sensitive to the possibility that his scheme may be taken to endorse libertarianism and goes to great lengths to argue that his scheme supports more than just a few presumptive goods. Yet even he admits that some desirable goods can't be justified on the grounds that they are somehow necessary for the later provision of presumptive goods. Some goods, such as funding for the arts (or maybe Nozick's public address system?), just look like things that many of us non-libertarians want but that can't be justified in terms of fair play.

Klosko's solution - to admit partial defeat and to say that other theories of obligation should pick up the slack here - is no solution at all. The “solution” seems to present Klosko with a dilemma. Either this move ignores the criticisms of these other theories (which are also the criticisms that make Klosko's scheme desirable in the first place) or it seems to make Klosko's theory redundant. For instance, consider the possibility that we supplement Klosko with a natural duty theory like that of Rawls (1971) or Waldron (1991).166 The idea here is that we have a duty to support just institutions that allow us to discharge our higher-order duties toward others. This theory has been criticized, most notably by Simmons167, on the grounds that such a duty can't easily tie me to a specific institution – a point I will discuss more.

167 In many, many places, including Simmons, *Moral Principles*, and Simmons, *Edge of Anarchy*, as well as his co-authored book Simmons and Wellman, *Duty to Obey.*
extensively in a later chapter.

My point here is not to debate the virtues of natural duty theory or any other of fair play theory's competitors. The point is this: if natural duty theory is broken, it's broken; appending it to another theory does not suddenly make it viable. If the supplemental theory is not broken, then surely a theory that can generate duties to support pure discretionary goods could also generate duties to support presumptive goods. If so, where does this leave fair play theory?

IV.3 Ambiguous empirical data

In the second appendix of his new edition of *The Principle of Fairness and Political Obligation*, Klosko argues that his fairness theory is buttressed by empirical evidence showing that people are only willing to take on obligations if they believe that burdens are distributed fairly. In fact, the attitudes run stronger still; people are willing to sacrifice in the service of a greater good if they believe that they're not being taken for suckers.

To support these claims, Klosko first cites evidence that shows that US taxpayers are more likely to cheat when they believe others are doing so. Later, he also provides statements by public officials claiming that they would be willing to have their constituents make sacrifices if other constituencies were similarly burdened. Klosko says that this shows that Simmons's characterization of how citizens view the state is wrong. The implication is that that citizens are neither clueless about the source of the goods they receive, nor do they deny that these goods come from
cooperative schemes. They also are ostensibly willing to pay more if necessary so long as other citizens also pay their fair share.

While some may say that using empirical data to support normative claims is mistaken, that is not the case here. That's because the commonly accepted metric by which theories of obligations are judged (in the current debate, at any rate) involves testing conclusions of a given theory with our considered judgments. If a logically impeccable theory says that we should push grandma down the stairs for fun, then we say that the conclusion is so at odds with our convictions that we feel entitled to reject it, logical impeccability and all. What's more, if people were explained fair play theory and asked if it applied, it would be a problem if they unanimously said that it did not. Hence Klosko is not wrong to invoke empirical considerations. Any theory of obligation must do this at some point. What's more, I do not contest these empirical data; in fact, Tyler (among others) has offered support for the thesis that Americans are much more willing to obey the law if they believe they are treated fairly.\footnote{See for instance Tyler, \textit{People Obey the Law}.}

The trouble with using this empirical data is that what respondents count as "fair" may turn out to be a type of fairness that Klosko rejects. Most of us think it proper that some of our relationships are hierarchical, and there's nothing that precludes this anti-egalitarian sense of fairness from being part of the political fabric. In a conservative society, people may agree that (a) fairness is important, but that (b) it is appropriate that different types of people hold different political stations, and (c) that some of these people should bear a burden that far exceeds that of others. When
we read Klosko's evidence, our egalitarian yearnings may cause us to want to interpret fairness according to egalitarian standards, but there are plenty of communities where it is understood that the fair thing to do is to distribute burdens unevenly. He thinks that a decision procedure’s being “democratic” is something that all would agree to, but is this necessarily the case? Without entering the morass of which democracies count as “real” democracies, what about those conservative societies where there seems to be a real popular support for regimes that are not formally democratic? Should we say that they are illegitimate? If so, it’s not clear on what basis we can do so. “Fairness” is culture and activity-specific. As Alan Fiske\(^{169}\) points out, all societies have relationships that people accept are differentiated and hierarchical. Fair play theorists like Klosko may need to accept that many members of some societies may well say “Yes, I am in favor of fairness. It is fair that the woman takes care of the domestic sphere while the man takes care of the political sphere.” His theory does not seem to have the resource to rebut such a claim.

### IV.4 Barely acceptable choices and symbolic violence

My final objection pertains to Klosko’s proposed system for choosing one set of policies over another. Recall the libertarian who argued that private entities could at least provide necessary discretionary goods. Klosko’s reply was that the libertarian was free to make her case, but that libertarianism was one package out of a set of defensible policy options, and so if another package from the set were chosen by a fair

---

\(^{169}\) In, among other places, Fiske, *Structures of Social Life*. 
procedure, the libertarian has a duty to go along with it. The same logic applies to the selection of selection procedures themselves: there exist a certain set of defensible selection procedures, and so long as the selected procedure provides presumptive goods and distributes burdens fairly, we have a duty to treat that procedure as legitimate.

But consider the following hypothetical case. A given decision procedure is in place and we are to select one of four possible option “packages”. Package A does not provide presumptive goods; Package B provides goods that, for a significant part of the population, barely satisfy Klosko’s conditions (i)-(iii); Package C is like Package B except for the fact that this “significant part of the population” is better off than in Package B. Package D is an even better version of Package C; everyone is doing great, although the best off are earning one dollar less. What are we to say if the decision procedure picks Package B? The decision procedure, we might think, was based on the principle of "one person one vote." What's more, that society - or rather the majority - has agreed that the preferred way of understanding democracy is to treat it as a system where all citizens are formally equal to use the state to pursue their material interests, so long as this use does not prevent anyone from formally exercising their democratic rights.

This, I believe, is a vision of democracy that many would not consider far-fetched or especially immoral. If I am a member of the aggrieved part of the population, do I accept the legitimacy of an existing system even though I know that Package D is possible and is only rejected because of the indifference or selfishness of
the winners? What objection can I give? Perhaps I can say that the decision procedure was unfair. But surely “one person one vote” and “everyone should pursue their private interests” is part of the set of acceptable choices. If it isn’t, Klosko’s reliance on intuition doesn’t tell us why we should reject this procedure. Is the substantive result unfair? I can, of course, make that case, but the majority might say that the result is defensible. They provide convincing (to them) reasons why the liberty that all citizens hypothetically enjoy is more important than the radical income disparity between groups.

As a final point, fair play theory is almost entirely silent on the question of symbolic violence. Consider the "separate but equal" policies that were in effect in the US for much of the 20th century. There are many reasons to object to these policies, and many of those reasons are very good. For instance, it’s doubtful that policies really could be separate and equal, both for theoretical and for empirical reasons. But even if it were possible, these policies would be evil and not worthy of respect because they constitute symbolic violence on a class of people.170 A fair play scheme that provides food is not worthy of respect if it forces some people to eat their impeccably portioned-out food out of a dog bowl. If the recipients of that food want to say that the scheme is illegitimate, I don’t see how fair play theory can help them. Goods are being provided, and they are being provided in ways that the majority sincerely thinks are fair. If symbolic violence does count, then I don't see how fair play theorists could tell (for instance) very religious groups, groups who feel that secularism is an affront,

170 The fact that it's a class of people isn't important here; the same point would apply if we were talking about particular individuals.
that they are to respect the law because they had a fair opportunity to participate. I will have much more to say about such symbolic offenses in later chapters.

V. Green and Reasonable Envy

While this chapter has been focused on exploring and challenging Klosko’s influential defense of fair play theory, I want to conclude by briefly discussing Jeffrey Green’s “plebian theory” of liberal democracy. His theory – expounded in The Eyes of the People (2011) and The Shadow of Unfairness (2016) – is not in direct dialogue with either Klosko or the philosophical anarchists, but provides additional challenges to fair play theory by questioning key assumptions in Klosko’s, and in many other democratic theorists’, work.

Green believes that we usually conceive of a well-functioning liberal democracy as one that identifies the voice of the people and translates it into policy. On this model, which he calls the “vocal model” of democracy, citizens understand themselves to have equal political rights, and advance their interests (if they so choose) by participating in politics. Against this view, Green claims that “mass representative democracy engenders and normalizes a type of citizen that, as a matter of law and abstract principle, has full political rights but, as a matter of practice, experiences politics primarily as a spectator.”\(^{171}\) Most citizens, Green argues, are interested in politics in that they consume political news, develop opinions about it,

---

171 Green, Eyes of the People, 32.
and so on, but do not actively participate in political life.\textsuperscript{172} As with elite theories, then, there are two effective classes of citizens: the few who are politically potent, and the many who are not.\textsuperscript{173} It is on this basis that Green believes that the “vocal model” of democracy, where citizens strive to make their voices heard, is flawed, and instead proposes an “ocular model”, where the citizenry constrains representatives – more akin to rulers in Green’s scheme – by subjecting them to their gaze and forcing them to perform in unscripted contexts.

If this is right, one problem from the perspective of legitimacy, and fair play in particular, is the existence of a gap between the promises of democracy and the lived experience of the citizens. Green identifies three structures that create the experience of second-class citizenship. First, citizens feel a sense of \textit{remove}. As most citizens mature from childhood to adulthood, they come to realize that they will never hold political office. What’s more, they also realize that they do not have enough access to office-holders for their opinions to matter.\textsuperscript{174} Second, most citizens come to learn that to the extent that they have any power, it is through \textit{manyness}, that is, through their conglomerating into a mass of people where the voice of almost any given individual

\vspace{1cm}
\begin{flushright}
\textsuperscript{172} Ibid., 33.
\textsuperscript{173} I do not want to give the impression that Green’s framework fits into an elitist theory of democracy, à la Schumpeter or Schmitt. While an involved discussion of Green’s differences with elite theorists would take me too far astray, the main idea is that whereas Schmitt proposes a plebiscitary model where a people give ascent to a powerful, unifying leader who articulates a national will, Green’s Weberian plebiscitary model acts as a check on leaders by forcing them to appear in public in unrehearsed contexts, and does not propose the existence of a national will that can be articulated. Against Schumpeter’s skeptical view of “the people” and of democracy, Green argues that “the people” do exist and can play a meaningful role in politics – just not the kind of role envisioned or hoped for by most democratic theorists. For more, see Green, \textit{Eyes of the People}, 166-177.
\textsuperscript{174} Green, \textit{Shadow of Unfairness}, 36-37.
\end{flushright}
is drowned out.\textsuperscript{175} Contrast this with the clear, articulate voices of powerful individuals or corporations that have access to political decision-makers. Third, we know that our voice is amplified and made more relevant by our financial wealth. To the extent that a society is \textit{plutocratic} – an inevitability in liberal democracies, even, Green argues, in states more egalitarian than the U.S. – we are cognizant of the fact that our effective political power does not resemble the frequently evoked vision of egalitarian citizenship.\textsuperscript{176}

A follower of Green might thus argue that fair play theory fails because it assumes the existence of a citizenry that views itself as politically equal and able to make its voice heard. Against this, the critic of fair play would go on to say that the inherent inegalitarianism of political life undermines exhortations that we all do “our fair share”. Such exhortations only make sense if one can make a plausible argument that (1) we are politically equal, a point that Green rejects, or (2) there is a good reason why we are politically unequal. But even if (2) is correct, there is no good reason to think that people who are political unequals should have the same duties. In fact, this is precisely the point that Green defends in much of \textit{The Shadow of Unfairness} and leads him to endorse apparently undemocratic ideals while maintaining a qualified endorsement for liberal democracy.\textsuperscript{177}

One such ideal is what he calls “reasonable envy.” Against the idea of viewing co-citizens as equal participants in a joint project, Green advocates singling out the

\textsuperscript{175} Ibid., 40.
\textsuperscript{176} Ibid., 44-46.
\textsuperscript{177} Ibid., 12-14.
super-rich and subjecting them to additional scrutiny and impositions. As he points out, the idea of singling out a class is not inherently controversial in democratic theory. Many theorists, including Rawls, have argued that the worst-off (however defined) deserve special and explicit consideration within a theory. There is no principled reason why this concern should be limited to the worst-off, and in fact a number of reasons to do so, even within a Rawlsian or generic liberal democratic framework.\(^{178}\) Green suggests that the rationale that allows us to consider one distinct class of people – the worst off – should allow us to consider another distinct class of people – the best-off. After all, if doing so allows us to propose a theory that better meets the needs of the worst off, how could we justify not doing so?

Green has two broad reasons for singling out the rich, and both depend on a class-centric view of liberal democracy. First, Green believes that we should recognize that the rich have the ability to violate civil liberties and shift effective political and economic rights in their favor.\(^{179}\) So far so good. But Green’s second argument, is more controversial. He argues that redressive measures are also appropriate in order to acknowledge and remedy, at least to some extent, the feelings of unfairness felt by the morass of second-class citizens.\(^{180}\) Green lists this as one way (but not the only way) for these citizens to come to terms with their relative impotence and maintain some level of political self-esteem. Claims that plutocracy emerges due to structural, not

---

178 Ibid., 77-80 & 98-100. Green’s defense of this position is complex and explaining it in detail would take me too far off course here.
179 Ibid., 82-83.
180 To be clear, this is not the only reason for proposing redressive measures. Green also discusses the types of reasons that might impress luck egalitarians, but since these types of reasons would not present a challenge to Klosko, I will not discuss them.
immoral, reasons do not impress Green. He acknowledges that the unfairness may not be deliberate, but this does not vitiate the lived experience of injustice felt by most citizens. What this means, and what Green also acknowledges, is that in his model, we are justified in imposing restrictions on the rich even when these restrictions do not materially benefit anyone. Not only are the poor and rich not equal participants in a joint venture, but the poor are entitled to highlight this inequality by hurting the rich.

The point is that reasonable envy, and Green’s theory in general, undermine the egalitarian logic of fair play theory. Green encourages us to think of the citizenry in a differentiated (in this case binary) way, with different citizens playing different roles, enjoying different benefits, and as a result holding different responsibilities. Rather than experiencing the state as a joint project, a normal citizen experiences it as a ruler. This does not mean that the citizen has no power. It does mean that the logic of fair play theories which presuppose a politically equal and potent citizenry fails to take hold in Green’s world.

VI. Conclusion

Although I have spent much of this chapter critiquing fair play theory, I would like to conclude by making clear the limits of this critique. First, as should be clear right now, much of the critique consists not of refutation, but of asking what I take to be difficult questions. Second, fair play theory is not nonsense. It is a powerful theory

181 Ibid., 9-10.
that generates clear obligations in certain cases. The question is whether or not it 
generates a clear general obligation to obey the law. Third, my emphasis on Klosko 
should be read for what it is: an emphasis on what I take to be the strongest proponent 
of fair play theory today. Klosko's scheme is a real achievement; his attempt to link 
 presumptive and discretionary goods is intriguing, but I believe that it does not 
succeed at clearly showing that a scheme that satisfies his conditions (i)-(iii) should be 
called legitimate. Some regimes may satisfy the conditions and still be evil. Fair play 
can’t justify the provision of a subset of goods. People might agree that regimes ought 
to be “fair” while advocating disturbing notions of fairness. Finally, it’s not clear how 
fair play deals with problems of symbolic violence and the selection of packages that 
are barely acceptable to a portion of the population. In sum, fair play theory should be 
considered suspect. If archists are right, they are likely better off making their case by 
appealing to a different theory of obligation.
Chapter 4: Voting and Consent

I. Introduction

Consent theory is one of the most venerable and persuasive theories of political obligation. Indeed, its appeal is such that even leading philosophical anarchists say that consent theory represents the most likely way that such obligations could be justified.\(^{182}\) Despite its appeals, most people who argue for the existence of general political obligations do not do so on the basis of consent. The fundamental problem with consent theory, the story goes, is that very few people consent to much of anything in the political realm.\(^{183}\) Theorists as far back as Locke have attempted to circumvent this problem by trying to ground political obligation on tacit consent, especially via residence. Powerful critiques of this position are nearly as old, going back to at least Hume, who launched a withering attack on consent theories of political obligation.\(^{184}\)

In this chapter, I take up the less explored argument that we consent when we vote. I argue that voting can generate consent in some cases, and so some voters may, by their act of voting, generate political obligations by way of consent. This is not

---

\(^{182}\) A.J. Simmons is quite upfront about this. In his seminal piece on political obligation, he says that consent *does* oblige, although it has trouble providing a type of obligation that would apply to all citizens. See Simmons, *Moral Principles*, 69-70. Robert Paul Wolff, the most prominent exponent of the other major branch of modern philosophical anarchism, says that the only legitimate form of government would be a unanimous direct democracy, where “... by the rule of unanimity a single negative vote defeats any motion.” Wolff, *Defense of Anarchism*, 24. His subsequent work, *The Autonomy of Reason*, is even more explicit: “I am persuaded that moral obligations [which include political obligations – see p.225], strictly so-called, arise from freely chosen contractual commitments between or among rational agents who have entered into some continuing and organized interaction with one another.” See Wolff, *Autonomy of Reason*, 219.

\(^{183}\) For one articulation of this position, see Klosko, *Principle of Fairness*, 142.

because of a necessary conceptual connection between voting and consent. Rather, voting and consent may be linked contingently; the social meaning of votes may differ, and whether or not a vote entails consent will vary accordingly. Voters may end up in a bind – they may want to vote for practical reasons while withholding their consent from the regime. I will argue that they can avoid their consent-derived political obligations only if they can signal the meaning of their vote to others.

This is important because modern philosophical anarchists, conscientious moralists that they are, may believe they have a practical duty to vote. Wolff and Simmons would be unimpressed by appeals to nationalism or civic duty (since they seem to deny that they are part of a political community in Strong's sense of the term), but they might think that they should nevertheless help elect the candidate that best represents their substantive moral positions. If a philosophical anarchist is pro-choice, for example, he might help elect candidates that, in a practical sense, help women have access to safe abortions. If voting implies consenting, however, philosophical anarchists who vote could find themselves doing more than helping bring about one set of policies.

This chapter is organized into two broad sections. In the first section, I unpack consent by considering its connections with promising. This analogy (and its limits) helps clarify the type of obligation inherent in consenting. In the second section, I reject the contention that voting is necessarily connected to consenting. I instead argue that the two concepts are often contingently linked. In such cases, voters end up tacitly consenting as a default when they cast their ballots.
Before beginning, however, I want to contrast this approach to my reading of Locke’s “tacit consent via residence” argument. I have two reasons for this. First, both Locke and I discuss “tacit consent”, but we use the term differently, even if neither of us uses it in an exotic way. Second, contrasting these approaches will help shine a spotlight on just what is at stake in the “voting as consent” model.

In his *Second Treatise*, Locke advances the following argument for the position that taking up residence in a location entails tacit consent. At this point in his overall argument, Locke assumes that we agree that express consent to enter into “society” (that is, the social compact that gives rise to government) is conceptually unproblematic.\footnote{Locke, *Second Treatise*, §119.} That said, what to make of people who have not expressed explicit consent? Locke is clear that the solution can’t be that the express consent of an individual applies to his descendants; I can’t bind someone merely by virtue of my consent.\footnote{Ibid.} So what do we make of someone of age who has not expressly consented to join a social compact?

Locke’s solution lies in his theory of tacit consent. Specifically, he says the following:

> And to this I say, that every man, that hath any possessions, or enjoynment, of any part of the dominions of any government, doth thereby give his *tacit consent*, and is as far forth obligated to obedience to the laws of that government, during such enjoynment, as any one under it; whether this his possession be of land, to him and his heirs for ever, or a lodging only for a week; or whether it be barely travelling freely on the highway; and in effect, it reaches as far as the very being of any one within the territories of that government.\footnote{Ibid.}

\footnote{\footnote{\footnote{\footnote{}}}}
In other words, living somewhere is one way that we consent to obey the laws of that place. It does not imply that we have entered into a compact. So long as I do not offer express consent to enter into a society, I am free to pick up and go somewhere else.\footnote{Ibid., §121.}

It is only when someone gives express consent to enter into a society that he “is perpetually and indispensably obligated to be, and remain unalterably a subject to it, and can never be again in the liberty of the state of nature,” barring an unforeseen calamity like government dissolution.

The main difference in the “voting as consent” argument that I will discuss in this chapter and Locke’s theory of tacit consent is that voting as consent avoids Locke’s position that duties differ depending on whether the consent given is express or tacit. In other words, if voting is consenting, then it expresses more than the kind of consent that Locke thinks applies when we are traveling through a country. Even if some does not believe that the act of voting itself means consenting, philosophical anarchists must take seriously the position that voting signals to others that one has already consented and is part of a political group, complete with duties that come with membership in said group.

II. Consent

The most straightforward way to categorize acts of consent is to treat the consenting act as express, tacit, or hypothetical. Roughly, express consent occurs when someone performs an action that serves as a signal of consent. Tacit consent
occurs when one can express dissent but remains silent, or who expresses consent incidentally by performing another action whose purpose is not inherently to express consent. To use A.J. Simmons' illustration, when a chairperson asks the attendees of a meeting if anyone objects to switching next week's meeting day from Thursday to Tuesday, their silence is just as binding as verbal assent. According to Simmons, both express and tacit consent will obtain as long as “[…] the means acceptable for indicating dissent must be reasonable and reasonably easily performed; and the consequences of dissent cannot be extremely detrimental to the potential consentor.” Although tacit consent may be harder to identify in practice, Simmons believes, contrary to Locke (at least with respect to Locke’s discussion in the Second Treatise, §119-125), that it nonetheless binds just as strongly as express consent.

Hypothetical consent, on the other hand, does not. Someone hypothetically consents when we believe that she should have consented to something had she had access to more facts or been true to her status as a rational being. To the extent that consent has a voluntarist character, and many liberals believe that it does, it cannot merely be hypothetical because hypothetical consent may obtain even if a person vociferously states that she is not consenting. If hypothetical consent binds, it does so for a different reason entirely and so falls outside the purview of consent theory.

189 Horton, Political Obligation, 29.
190 Simmons, “Tacit Consent,” 279-80.
191 Simmons, Edge of Anarchy, 81.
192 For the seminal two-part piece on hypothetical consent, see Pitkin, “Obligation and Consent I”; and Pitkin, “Obligation and Consent II.”
193 Waldron, “Theoretical Foundations.”
These rough definitions are too rough, however. After all, if I make a gesture that signals express consent, what does mean? We need to know something more about consent before we can use it in the definition of express or tacit consent. The rough definitions are good enough given an unreflective but usually serviceable understanding of consent, but they are of little help in hard cases where our intuitions make us anxious about definitively classifying a given act as consenting-signaling. The remainder of this section will discuss consent in greater detail by considering its connections with promising. Doing so will provide the conceptual apparatus that I will use in the next section when I discuss the connection between voting and consent.

II.1 Affinities between consent and promising

Consent shares many important features with promising, and since we have a decent intuitive grasp on what it means to make a promise, drawing an analogy between the two will take us quite far. Indeed, some theorists feel comfortable using the terms synonymously. There are at least three affinities between consent and promising. First, like promising, consenting is a discrete voluntary act that generates obligations of some kind. These obligations may be to perform an action, to refrain

---

194 As Dworkin famously put it, "[...] hypothetical contracts do not supply an independent argument for the fairness of enforcing their terms. A hypothetical contract is not simply a pale form of an actual contract; it is no contract at all." Dworkin, “Original Position,” 17-18.

195 Murphy, Natural Law, 93.

196 For more on this, see again Waldron, Theoretical Foundations; and Simmons, Edge of Anarchy, chapter 3.
from performing an action, or some combination of the two.\textsuperscript{197} If I consent to meet you at the bar this evening and I don’t show up, most people would accept that you would have a prima facie right to reproach me by reminding me of a precise moment in time when I performed what we both considered to be an obligation-creating action.\textsuperscript{198}

This is important for those of us who are concerned with our status as autonomous beings. As Simmons says, “First, […] Consent theory maximizes protection of the individual's freedom to choose where his political allegiance will lie. […] Second, the model of the promise lends clarity and credibility to a theory of political obligation: for promising is surely as close to being an indisputable ground of moral requirement as anything is.”\textsuperscript{199} Contrast this with other ostensible bases of political obligation, such as appeals to fairness.\textsuperscript{200} Such appeals have their strengths but suffer from a comparative lack of clarity (“Do I actually have an obligation in this case? What exactly is my obligation?”) and voluntariness (“Did I incur an obligation without intending to do so?”). From the perspective of those concerned with autonomy, there is something disconcerting about obligations that arise independently of one's voluntary actions.\textsuperscript{201}

\begin{itemize}
\item \textsuperscript{197} Hyams, “Consent Doesn’t Work,” 111.
\item \textsuperscript{198} These include speech acts such as those discussed in Austin, \textit{Things With Words}, 5. I discuss these in greater detail in Chapter 8.
\item \textsuperscript{199} Simmons, \textit{Edge of Anarchy}, 71. Note however that Simmons agrees that promising does not equate consent. For his discussion on the matter, see 75-78.
\item \textsuperscript{200} For a relatively recent version of this type of appeal, see Klosko, \textit{The Principle of Fairness and Political Obligation}.
\item \textsuperscript{201} To be clear: I am not saying that all obligations do or should arise from voluntary actions. I am saying that the apparent clarity of these obligations makes them more appealing, at least to philosophical anarchists.
\end{itemize}
Second, promising and consent seem to generate obligations in roughly the same way. To see this, consider our reactions when someone breaks a promise. We may react in any number of ways, but the paradigmatic feeling is one of disappointment. One could ask why disappointing others should be a moral wrong. The answer is that promising – or consenting – set up a particular kind of disappointment. When I promise something, I create new obligations for myself, and in so doing generate claim rights for the promisee. These rights give moral force to the disappointment. It is one thing when I feel disappointed that my friend will not quit smoking, but it is quite another when I feel disappointed that my friend keeps smoking after promising that he would quit immediately. Even if I have no prior right to demand such a promise, once it is made I now have a right to make particular demands of my friend.

This disappointment is connected with our having anticipated future events. The consent theory literature frequently couches the wrongness of violations of consent-generated rights in terms of practical harm. For instance, Albert Weale says that

By inducing reliance upon one's future performance (or non-performance), one is encouraging others to undertake their plans and commit their resources to a future course of conduct which they have reasonable grounds to believe will be supported (at least by non-interference) by the speaker. The speaker is, therefore, obligated to act in accordance his consent by the general utilitarian principle that where others have planned their own future conduct relying on an undertaking one has voluntarily given, then one is obligated not to act in violation of the undertaking in question.\(^{202}\)

\(^{202}\) Weale, “Consent,” 69.
While this is a reasonable interpretation of disappointment from a utilitarian framework, one need not be a utilitarian to condemn broken promises or violated consent. Promises and consent may also signal changes in relationships that have no immediate practical consequences. The symbolic assertion of a particular type of connection may be important in itself. In other words, someone may be morally wronged by the violation of a promise even if the breach in no way affects what that person can or cannot do.

For instance, my promise to be my niece's godfather goes beyond signaling to her and her parents that they can expect certain types of actions. I may consent to the role even if all parties know that I have a fatal disease that limits my life expectancy to a few bedridden years. From the point of view of my niece, the importance of the relationship created out of my consent may outlast my life. This type of reasoning can be – and as I shall argue in the next section of this paper, should be – extended to the wider political realm. As Joseph Raz puts it, "There are various attitudes toward society that consent to the authority of law can express. They can all be regarded as so many variations of a basic attitude of identification with the society, an attitude of belonging and of sharing in its collective life." This additional understanding of the harm caused by violating promises or consent will, I will later argue, complicate the link between voting and consent.

---

203 I should point out that some utilitarians might object to any instance of consent. For one prominent example, see Godwin, Concerning Political Justice, 150-56.
204 Raz, “Authority and Consent,” 125.
205 I will parse out the difference between promising to be the godfather and my consenting to do so in the next part of this section.
206 Raz, “Authority and Consent,” 128.
Finally, promises and ostensible acts of consent share similar restrictions on their capacity to bind. Many of these boundaries will be familiar to most political theorists and have been neatly summarized by Keith Hyams. According to Hyams, an attempt to consent fails if any of the following obtain: (1) an individual's alternatives are sufficiently unpalatable, especially if this is the result of coercion, (2) an individual does not understand that the “consenting” action in question counts as consenting, or has false beliefs regarding the background and consequences of this consent, (3) an individual does not have the general competence to consent, that is to say that the individual's inability to consent is not situation-specific. And (4) an individual is not authorized (morally or otherwise) to consent, given the substantive particulars of the agreement.\textsuperscript{207} These limits apply to both express and tacit consent and serve as a reminder that these two "types" of consent are and as such help us treat them as instances of the same underlying phenomenon.\textsuperscript{208} These criteria apply equally well to promising. If we know what it means to make a promise, it would seem that we know a great deal what it means to consent.

II.2 Differences between consent and promising

I have argued that promising and consent are analogous in that they (1) are voluntary obligation-creating acts that (2) create obligations on the promisor and rights for the promisee and that (3) are subject to roughly the same formal limits with respect

\textsuperscript{207} Hyams, “Consent Doesn't Work,” 112-119.
\textsuperscript{208} This is useful because relying too heavily on the distinction between express and tacit consent may lead us to erroneously believe that these two types of consent are different in a fundamentally important way.
to their applicability and scope. Yet exploring the disanalogies between promising and consenting will prove to be useful in a later discussion of the obligatory force of voting.

First, consenting must involve multiple people. This need not be the case with promising. Looking down at my chocolate-smeread hands, I can tearfully cry out that I promised myself that I wouldn't eat the whole cake, but, assuming the cake was mine to eat as I saw fit, I could not lament that I consented to myself not eat the cake. Another way of putting it is that while promising requires two people (since X promises Y that she will Z), a promise can occur when both people have the same identity (e.g., X promises X that she will Z) or when the promisee is hypothetical. Consent, on the other hand, can only happen between actual and different people. It makes sense to say that I promise (to) myself that I won't eat the cake, but telling myself "I consent not to eat the cake" implies a separation of self that seems too extreme. I can promise my deceased grandmother that I will quit smoking, but I cannot say that yesterday I went to her grave and consented to stop smoking. Voting, I will argue, is an act that is necessarily other-regarding.

The fact that consent is other-regarding can be seen by looking at another example of self-promising, one that does not raise questions of my ability to follow through, as with the cake example. Imagine I, in an effort to do a good turn, set off to help my neighbor with her home repairs. I soon find out, however, that she is a very unpleasant person. Now, it seems possible to imagine a scenario where I promised her that I would help. In other cases, I could have consented to her request that I help. In
either of those cases, it would be difficult to extirpate myself from the engagement. But I also could have promised *myself* that I would help her. This is a personal commitment, one that seems far easier to break when I find out just how unpleasant she is. It seems far easier to tell myself “I’ll just help my other neighbor instead” than to tell her that the deal’s off. Note, however, that I cannot consent to myself to help her. Again, the separation of self is too vast here to pull off that linguistic trick.

If promises are broader than consent in one respect, they are narrower in another. Promises are always either future-regarding or reality-affirming. The most typical type of promise is one that commits me to do something or to refrain from taking action at some point in the future. For instance, a promise to be my niece's godfather implies that I must take on that role at some appropriate time in the future. I can also use promises to reassure someone that we do in fact have a certain type of relationship. Someone may tell an insecure friend “I know that you worry that I take you for granted, but I promise you that you're my very best friend.”

Consent may also be future-oriented or reality-affirming, but it has the added feature of being able to generate a new type of relationship in the present. I may promise my sister that I will be my niece's godfather, but at a given moment in time I must *consent* to take on the role. My promise may commit me to consent at a future date, but my status with respect to my niece only changes once I have actually consented. Once I have done so, my relationship with my niece changes immediately. The same is true when I agree to waive certain rights. A promise to waive rights is a commitment to do so in the future; consenting to waive rights may in some cases be
similar to a promise, but in other cases, my consent constitutes the precise moment in which the rights are waived.

That promises do not create immediate new relationships might be best understood by considering that promises are usually only made between people that already know each other. Promises are typically informal. It would be odd to make a promise to a stranger without at least an indirect appeal to a shared identity (e.g., "I promise you, from one gentleman to another, that I will pay you back.") or to the prospect of likely future encounters (e.g., promising the manager of a bar I frequent that I will pay my tab the next time I come in for a drink). Promises, in other words, are personal, and in such cases there simply is usually no cause to establish a fundamentally new type of relationship between people. Even in the context of a formal setting, a promise is a type of aside – when I promise the judge that I will change, I am interjecting a moment of personal appeal in an otherwise impersonal setting.

Consent, on the other hand, is more formal than promising and adds an impersonal quality or formality to a relationship. This is not to say that consent cannot take place between intimates, but it does so in a peculiar way. Consider once again the

---

209 Tracy Strong has pointed out to me that swearing allegiance appears to be a counterexample. This is a good point, to which I offer the following responses. First, I am using the terms in their ordinary sense. This response amounts to punting and accepting that there are exceptions to the distinction I draw. Second, it’s worth noting that most people probably wouldn’t call “swearing allegiance” promising if it doesn’t contain the actual language of promising. There is a reason why we reserve the expression “swearing allegiance” for rare occasions, after all. Third, if the swearing of allegiance does contain the language of promising, it doesn’t necessarily mean that promising is taking place. This type of issue occurs frequently with respect to the language of apology, which I discuss in a later chapter. Empathizing with a friend by saying “I am sorry” is not a kind of apology, even if it contains elements of apology (i.e., empathizing and recognizing both her pain as well as the possible unfairness of her situation).
case of my consenting to be a godfather. Although I know my sister and my niece, my consent marks a special, ritualized occasion in which others recognize a sudden shift in my formal relationship with my niece. Or consider consent's role in sexual encounters. Consent's formal connotations are clear in cases where strangers engage in anonymous sex. Consent, in this case, refers to and introduces a codified, institutionalized (by reference to law, morality, or both) interpretation of the relationship. This is obvious in the case of strangers, but it also applies to long-standing, emotionally intimate partners. The fact that consent occurs, even if it is unstated (say in the context of a monogamous couple that has long-standing mutually desired practices of sexual contact), means that the partners' status towards each other is to some extent a formal one. Consent, in other words, simultaneously binds and distances us.

II.3 Consent, promising, and informality

Consent is thus what allows us to make promise-like agreements with strangers. Knowing this, we may be tempted to come to the erroneous conclusion that the informal limits of promises are socially determined and that the boundaries of consent are explicitly specified. This overstates the difference between the two. All instances of promising and consent, however formal or informal, rest on unspoken and in some cases possibly unspeakable understandings of the content and limits of the

---

210 To the extent that I can “know” a baby.
211 This formality does not need to be encoded in a system of laws if the social understanding of the impact of my decision is sufficiently well-established.
symbolic actions in question. This is evident in the case of promises. To illustrate: a friend may ask me as we enter a bar, "If anything goes down, do you promise that you have my back?" If I say "yes," a promise has been made even if neither of us is precisely sure what exactly is meant by "goes down" or to "have someone's back." By saying "yes," I clearly agree to perform some types of actions by virtue of my status as a friend. Given our relationship, this may mean that I should talk him up if he starts flirting with someone. Yet my friend cannot take my promise to mean that I will fight alongside him if he decides to start a brawl. This is because he is my friend and he knows that I would not agree to such a commitment. The promise would not be binding because it would violate Hyams' conditions; I would not know what the promise meant.

In some cases, and with some people, however, "having someone's back" has nothing to do with helping them flirt and really does mean to fight alongside them if a brawl erupts for any reason. In each case, the content of the agreement is filled in after the agreement has been made, and the content is limited to our mutual understanding of what types of practices agreement could entail in this context. To the extent that we share an understanding of the local rules, agreements need not be spelled out for obligations to be created. Elaborate specification is simply unnecessary

When two people are acutely aware of their different interpretive strategies, we move from the realm of promising to consenting, and this requires more explicit agreements.

---

212 This is in line with Pateman's assertion that “[…] if citizens are to consent then, as in the social practice of promising, they must be able to ascertain what kind of commitment they are undertaking and whether good reasons exist for them to do so.” Pateman, Political Obligation, 88. The key is the specifics need not (and cannot) be fully fleshed out prior to the act of consent.
Even cases of consent between strangers rest on unspoken understandings. In some cases, the participants may know the boundaries and meaning of consent without ever having articulated them. Consider the following example. Jill and I play speed chess for twenty dollars a game at a cafe. After getting thrashed repeatedly, I decide to call it a day and leave. Dejected, I silently place the chess pieces back on their starting squares. Jill, sitting in front of the white pieces, immediately moves a pawn up two squares, starting a new game. In this case, my intended meaning was to set up the board for the next player, but Jill could plausibly say that she understood my action as an indication that I was consenting to play her again. Indeed, that was part of my intention the previous times I set up the board. If I were to get up after this most recent pawn move, Jill would earnestly say that in doing so I would be forfeiting the game and that I would owe her another twenty dollars. By setting up the pieces, she would continue, I made a series of active gestures that are similar in kind to yelling “aye”.

This example shows that intent matters because it allows others to assess my competence. There are three questions to be answered here. First, was I being coerced into taking action? Second, how firmly ingrained were the social expectations attached to my action? Third, and by way of conclusion: am I competent enough to consent? In this case, the answer to the first question is obvious: there is no coercion.

Answering the second question requires some knowledge of on-the-ground
facts. What does my setting up a chessboard mean? In most cafes, it would likely be understood as expressing interest in starting a new game, but not that I grant the other person the right to take my twenty dollars if they start moving pieces without some other signal of assent on my part. But in some cafes this is not the case; in these places “everybody knows” what setting up the board means. If I go to such a cafe for the first time and am surprised by this unspoken rule, my consent may be vitiated on the grounds that I am incompetent within that context.

The answer to the third question will depend on the answer to the second one. My incompetence may vitiate my signal of consent, although it would do so at the cost of marking me as an outsider, one who likely would not have access to some of the formal or informal privileges of membership in the group. Note that the likelihood of my consent's being vitiated is tied to how reasonable others believe my ignorance to be. To the extent that others see my ignorance as unjustifiable, I am liable to pay heavier and heavier social penalties for my social incompetence. I cannot eat my cake and have it too; I cannot excuse myself from my obligations while maintaining the social and psychological advantages associated with a connection to a meaningful community. It is useful to pause and remember that at this point, my incompetence is purely contextual – no one doubts that I am a sane man who is capable of consenting in many other contexts. It is only if I were to repeatedly violate these norms that I would be deemed either morally blameworthy or incompetent in a general sense. Finally, so long as others see my incompetence as contextual, they may come to

214 Note that this “meaningful community” is not necessarily the same thing as a political community in Strong’s sense.
justifiably argue that my ignorance does not excuse the fact that my apparent consent created legitimate expectations on their end and that I am bound to follow through as though I had wanted to consent. It is only when there is a consensus that my incompetence is far-reaching that this line of argument fails. In such a case, and only in such a case, others would judge Jill to be in the wrong when she demands payment for the aborted game. This scene, I will argue, is analogous to some cases of voting.

III. Voting

If we begin with the premise that voting is connected to obligation, we can easily see the affinities between voting and consent. First, voting does not in principle violate any of Hyams' four restrictions on the scope of consent. Second, voting is a voluntary, symbolic, other-regarding act. Yet voting is not necessarily akin to promising, although it may be in some cases. Most instances of voting are tied to an institutional structure and take place without regard to the personal relationships one may have with other voters. For a vote to be a promise, the voter would have to have quite a high affinity with the other voters.

Voting may bind in at least two different ways, and only one of these is compatible with the practice of promising. On the one hand, we may see a vote as a commitment to perform certain actions in the future, such as obeying the law or

215 I will return to this premise later in this paper.
216 There are particular cases where this is not true. For example, voting can violate the non-coercion clause. Even well-regarded democracies such as Australia violate this clause when they fine citizens for not voting. My analysis does not apply to such cases. To put it bluntly: voting does not involve consent in Australia.
refraining from overthrowing the elected government. This interpretation describes voting as a future-regarding act and is consistent with interpreting votes as either promising and consenting. On the other hand, recall Raz's assertion that consent can also create relationships. Voting may do the same. As I will discuss below, the social significance of the vote may in some cases be analogous to my agreeing to be a godfather. It has more to do with the creation of a meaningful identity than as a commitment to perform concrete actions.

III.1 Voting does not entail consent

Before discussing how voting can bind, it is worth considering and dismissing the argument that voting entails or is constituted by consent. J.P. Plamenatz takes this position in his postscript to the second edition of *Consent, Freedom, and Political Obligation*:

Surely, to put a cross against the name Smith on a ballot paper at an election is to express the wish that Smith should hold the office for which he is a candidate? But the voter may not want Smith to hold the office; he may think that Smith has no chance of getting elected and may vote for him only for the sake of preventing the election of Brown. Yet, if Smith were in fact elected, it would be odd to say of anyone who had voted for him that he did not consent to holding the office.  

The idea expressed here and throughout the postscript is that so long as elections are free and fair, to vote means to “… consent to the authority of whoever is elected to the office.” In other words, Plamenatz is saying that it is correct but insufficient to say that the symbolic meaning of my vote for Smith is merely that I consent to give my

---

218 Ibid., 170.
vote to Smith. It is insufficient because “giving my vote to Smith” can only be done in a context where it is understood that the meaning of my so doing is tied to my participation in a scheme whereby authority is (usually) conferred to the person or party that gets the most votes. Voting for Smith, the argument goes, cannot even be a coherent act without this background, and so to give him one’s vote necessarily means to endorse the procedure that makes voting conceptually coherent.

This is too strong, however, and this position is subject to a number of replies. First, many people who repudiate the democratic process vote nonetheless. A subset of revolutionary socialist groups in particular has advocated using voting (among other methods) to lay the groundwork for a socialist revolution that would overturn liberal democracy. Lenin, chastising the German Left for considering parliaments obsolete, summarizes a point by saying that

> The inference which follows from this is quite clear: it has been proved that participation in bourgeois-democratic parliaments a few weeks before the victory of the Soviet Republic, and even after that victory, not only has not harmed the revolutionary proletariat, but has actually made it easier to prove to the backward masses why such parliaments should be dispersed, has made it easier to disperse them, and has facilitated the process whereby bourgeois parliaments are actually made “politically outworn.”

Revolutionary socialists who vote are not being incoherent; they are using means at their disposal to achieve their goals.

One objection to this may be that these revolutionary socialists’ actions are parasitic on the understanding that voting implies consent. If everyone acted as they

219 Lenin, *Left Wing Communism.*
did, the objection continues, there would be no voting procedures at all. In other
words, this objection essentially says that revolutionary socialists who vote are as
coherent as liars. The fact that a liar is acting coherently insofar as she consistently
acts towards achieving an aim does not excuse her from the obligations she incurs
when she utters statements that others take to be true. Analogously, perhaps the
revolutionary voter incurs obligations regardless of why she votes, since the meaning
of voting cannot be understood without reference to binding consent.

But there is an important distinction here that differentiates liars from
revolutionary voters. Liars implicitly endorse a persistent regime of truth-telling so
that they can maximize the benefits of their lies. The longer the regime of truth telling
holds, the better for the liar. This is not so for the revolutionary socialist who votes.
This type of voter openly wants to remove the regime of representative democracy and
does not necessarily object to the entire process becoming degraded. What these voters
are doing is tantamount to a liar announcing “I don't owe you the truth” before making
a statement.

The second reply is that in some accounts of democracy, the values that justify
democratic procedures in some cases may lead them to undermine democracy and
voting in others. Consider the attitudes of some citizens living in what Ronald Dahl
called a "polyarchy." Polyarchy, for Dahl, is characterized by a system of
representative voting where each citizen's vote has equal weight, where they can vote
freely, where the candidate with a plurality of votes wins, and where elections take
place regularly. The citizens might find this type of regime legitimate because they judge it to be the best chance of furthering their substantive interests without risking a radical reversal of fortune. These people advance these interests by joining groups, pressuring governments, forming coalitions, and so on. Nothing in this story says that anyone sees democracy as something inherently valuable.

Given the antecedent values that justify democracy in something like a polyarchy, to vote in this case does not seem to give any deep legitimacy to the democratic procedures as such, and the superficiality of these democratic procedures will presumably be well known. For example, if the antecedent value is negative liberty, then a given person's vote cannot confer deep legitimacy to a regime which, it is later discovered, turns out to be suboptimal in terms of promoting that antecedent value. The same holds true if the accepted antecedent value is egoistic enjoyment. This is again different from the case of the liar. If most people sincerely endorse an instrumentalist version of democracy like the one I just described, then the truth of the symbolic statement of each act of voting must be understood not as a binding social norm, but as an instance in a series of calculations in a kind of game. To deceive others in such circumstances is clever sportsmanship, not a moral violation.

Third, it is not clear what, other than Smith's being in office, Plamenatz' voter is actually consenting to. Governments are complex entities; even the simplest ones

---

220 Dahl, Democratic Theory, 84. I have listed the main salient features; Dahl’s formal definition contains additional clauses.

221 As Rex Martin argues, no regime that relies on an external standard in order to generate legitimacy is guaranteed to be safe from being replaced by a regime that is taken to be more in line with that standard. See Martin, “Wolff's Defense,” 143-47.
have sub-units, units which are frequently themselves organized via elections. If Smith is running for a municipal position, I understand Plamenatz to be saying that I am consenting not only to Smith's winning that municipal election but also to the legitimacy of the municipal election as such. But municipal elections are held under the auspices and rules of state or provincial governments; as such, the rules for municipal elections can usually be changed by these higher-order governments. Does my vote for Smith also constitute consent to this oversight and set of rules as well? If state or provincial governments are under the authority of a federal government or a body such as a supreme court, must my vote for Smith for town sheriff imply a far-ranging endorsement of the national government? We know that this can't be the case, but the conceptual argument that voting necessarily implies consent has difficulty specifying the limits of that consent. As I will argue, this problem dissipates if we accept the contingent link between voting and consent; in such cases, contingent factors determine the substance and limits of consent.

Two further examples will hopefully illustrate that voting is not necessarily attached to consent. First, governments sometimes hold non-binding referendums – plebiscites – to gauge public opinion or to help legitimize a controversial course of action. Certainly, voters in these plebiscites are not consenting to anything when casting their vote. Perhaps someone might reply to this that plebiscites are not cases of actual voting and that they amount to glorified opinion polls. Real voting, this line of thought continues, must imply a symbolic action that directly contributes to a decision. This reply is unsatisfactory. If the reply is saying that voting is constituted by consent...
because voting must imply making a decision, then it comes dangerously close to question-begging. If instead the reply is trying to show that our ordinary use of the term "voting" implies consent, then it must contend with the retort that people who take part in plebiscites do call what they are doing "voting."

Finally, people vote in autocratic regimes, and it is not clear what to make of the meaning of these votes. Few people seriously believe that voting in such cases constitutes taking on any real obligation with respect to that regime. The point isn't that voting in such conditions violates Hyams' non-coercion clause; in many cases, citizens are free to abstain from voting. Indeed, many autocratic regimes want nothing more than for their citizens to abstain entirely from politics. The point is that if voting implies consent, then voting in such a regime would imply consenting to obey an unjust regime.

Consent theorists may disagree among themselves about why voting in this case would not constitute real consent – some would say that it's because one cannot consent to the performance of an unjust act while others would say that we could not plausibly expect that people would actually endorse this kind of regime – but virtually all would agree that this type of voting should not bind.

Nor is the problem with voting in autocratic regimes the fact that one’s vote is usually meaningless. A meaningful vote between two unjust options should also not bind.222 In some cases, the lack of just options is obvious to everyone. In other cases,  

222 This was nicely illustrated in an early episode of The Simpsons, “Treehouse of Horrors VII.” Two hostile space aliens, Kang and Kodos, working in tandem, manage to make themselves the presidential candidates for the Democratic and Republican parties, thus becoming the only viable choices. In the final scene, humanity is enslaved under President Kang. When Marge complains,
such as in pseudo-democracies, the justice of the regime is ambiguous and subject to reasonable disagreement. These are not cases where voting must imply consent. When there are multiple bad options, someone may sensibly vote for the slightly lesser of two evils while denying that the regime is legitimate or that his vote constitutes consent to anything whatsoever.

III.2 Social understandings of the link between consent and voting

The upshot of this extended discussion is that Plamenatz’ argument that voting and consent are conceptually linked is mistaken. I instead take the more modest position that if voting is linked to consent, it is because there is a contingent but socially accepted link between the two. This is a position many people seem to hold intuitively, and has a historical basis. For instance, Bernard Manin argues that governance by lot disappeared with the rise of government grounded on the principle of consent, which in turn was to be expressed via voting:

“Under an elective system […] the consent of the people is constantly reiterated. Not only do the people agree to the selection method – when they decide to use elections – but they also consent to each particular outcome – when they elect. If the goal is to found power and political obligation on consent, then obviously elections are a much safer method than lot.”

The importance of this consent grew from the Middle Ages onward as repeated

Homer replies "Don't blame me! I voted for Kodos!" Groening, Eighth Season.
223 Indeed, in such cases a common strategy by leaders of such regimes is to encourage participation in elections in order to draw on the legitimacy that voting confers in robustly democratic regimes, all while trying to manipulate elections just enough to block regime change. For example, see Schedler, “Nested Game.” For further disambiguation on why different types of authoritarian regimes have different goals qua citizen participation, see Linz and Stepan, Democratic Transition, 44-45.
224 Manin, Representative Government, 85.
application cemented a belief that consent was a source of political legitimacy, and representative government emerged as "... citizens were now viewed primarily as the source of political legitimacy." Voting, consent, and legitimacy seem intertwined as a matter of fact even if neither concept logically entails the other.

This contingent understanding of the link between consent and voting is useful in helping us determine what is going on when someone believes that she is consenting when she casts a vote. As I pointed out in my discussion of Plamenatz, even if he is right when he says that voting confers legitimacy by way of consent, he cannot clearly state the precise object of this consent. Achim Hurrelman et al. argue, for instance, that voting can confer legitimacy to at least four conceptually separate levels of government. So far as I can see, there is no way to specify a priori to which level(s) are being legitimated when one casts a vote.

This problem, while conceptually insoluble, is as a practical matter frequently easy to sort out. Take for instance a relatively hands-off dictatorial regime, one that holds sham national elections. In this case, it is possible that "everybody knows" that voting in the national election is something of a joke, but that voting in certain smaller scale, local elections is meaningful. It may be widely accepted that voting in these smaller elections does imply consenting to the local authorities or decisions pertaining to specific governance issues without implying that one is consenting to the national regime that serves as a backdrop for these local elections. Voting, like playing chess in

225 Ibid., 88-92.
226 Specifically, and from most to least fundamental, voting can indicate consent to (1) regime principles, (2) the political system or community as such, (3) core political institutions, and/or (4) groups of actors. See Hurrelmann et al., “Democratic Nation-State”, 495-499.
a cafe, has a very clear meaning to the participants, and this meaning may be difficult to see from the outside looking in.

Without trying to catalog an exact list of obligations we incur when participating in elections, what kinds of things do we agree to when we vote? In some cases, it might be understood that voting means performing specific actions that support the regime or refraining from taking actions which harm it. Some such actions may be quite clear given the context of the election. Yet voting often goes beyond a commitment to particular actions and may also constitute a signal that the voter is a member of the democratic polity in some meaningful sense. Just as how my agreeing to be a godfather means something that goes beyond the pragmatic, in some cases being a voter might signal to other members of the community that I am entering into a symbolically significant relationship with other voters.

This relationship does not necessarily imply the creation of any practical duties, and yet finding out that one was deceived about the nature of such a relationship may sting nonetheless because this relationship helps constitute our identity. To many people, seeing someone with an “I Voted!” sticker pasted on their shirt implies more than a belief that the voter committed to abiding by the results of a collective decision-making procedure – it implies that the voter feels a sense of belonging or patriotism that is borne from participation in a process. From this perspective, the vote either creates a shared sense of identity or re-affirms it, a bit like a political equivalent of renewing one’s wedding vows. I speculate that this partially explains the anger that many people feel towards apathetic non-voters. These non-
voters appear as people who are explicitly rejecting a meaningful relationship with her co-citizens.

Michael Stoker presents a classic example that can help us get a feel for the intuition at play here. He presents the following thought experiment:

But now, suppose you are in a hospital, recovering from a long illness. You are very bored and restless and at loose ends when Smith comes in once again. You are now convinced more than ever that he is a fine fellow and a real friend-taking so much time to cheer you up, traveling all the way across town, and so on. You are so effusive with your praise and thanks that he protests that he always tries to do what he thinks is his duty, what he thinks will be best. You at first think he is engaging in a polite form of self-deprecation, relieving the moral burden. But the more you two speak, the more clear it becomes that he was telling the literal truth: that it is not essentially because of you that he came to see you, not because you are friends, but because he thought it his duty, perhaps as a fellow Christian or Communist or whatever, or simply because he knows of no one more in need of cheering up and no one easier to cheer up.²²⁷

If most people believe by convention that voting implies the creation of a certain kind of relationship with other voters, then if I vote while inwardly denying the implications of my vote I harm those who reasonably understand my vote as more than practical gesture to try to bring about one set of outcomes over another. It is one thing if I merely refuse to consent; it is far worse if I effectively tell someone that we have one kind of relationship only to have her discover after the fact that I had lied.

The less a vote's practical value, the more it becomes an exclusively symbolic act that can bind. On the surface, it would seem that active involvement in the

²²⁷ Stoker, “Modern Ethical Theories,” 462.
democratic process, including the possibility that one's actions could create a meaningful change in the outcome, should bind citizens more strongly than a throw-away vote where they are one of several million. This may be the case for reasons pertaining to other theories of legitimation, but from the perspective of consent theory it is precisely when my action can have a direct impact on the world that I can plausibly argue that the purpose of the action was this practical effect and not a signal of consent.\textsuperscript{228} In other words, if an action (such as casting a vote) can be understood to have concrete effects, we may argue that these effects – not consent – constitute the purpose of the action. “Meaningless” actions have no concrete effects, and so can only be plausibly understood in terms of their signaling consent.\textsuperscript{229}

\textbf{III.3 The non-consenting voter}

This does not bode well for someone who wants to vote for practical reasons but who does not want her vote to imply consent. This voter may be able to get away with voting in small, local elections, on the grounds that her vote actually might make a difference. She could not plausibly say that this is the case, however, when she decides to vote in the next presidential elections. In the vast majority of cases, she will not believe that her vote will be decisive (or even that her vote will be the decisive example that incites others to vote), and others will know this as well. In many cases,

\textsuperscript{228} Again, I should emphasize that it is not the mere fact of my intentionality that matters, but the fact that others may plausibly believe that I intended something other than consent. In other words, the nature of my action is, from the perspective of others, indeterminate.

\textsuperscript{229} This also applies to actions that people believe to be meaningless. The symbolic meaning of my vote does not change even if it turns out that my vote turns out to be pivotal.
people will thus understand the vote mainly in terms of its expressive, symbolic value.

Can this non-consenting voter vote in good conscience? The answer must be "it depends." To the extent that voting is socially understood as a practical act, she can do so. In such a case it would be the few who believe that voting should have an additional symbolic import who would be frustrated come election time, since they would be the eccentrics who see elections as more than a mere exercise of power. But to the extent that voting is seen to imply that one is performing an act that either creates political obligations to act or demonstrates allegiance to a meaningful community, the non-consenting voter has a problem. Even in cases where the social significance of voting has been brought about for cynical reasons – e.g., propaganda, jingoistic rhetoric, etc. – it nonetheless has a meaning that others take seriously and that is understood to constitute consent.

To return to Locke’s distinction that I mentioned in the introduction, we can say that the voter does not get to decide whether or not her vote makes her a member of a political community in the same way as would taking an oath. Intentionality matters, but not in the way that some would-be voters may think. Intentionality could refer to deliberately seeking an outcome or could refer to deliberately undertaking actions that we know will lead to an outcome. The former is surprisingly irrelevant in this case. We don’t get to dictate the consequences of engaging in social practices that are, strictly speaking, unnecessary to our living a reasonable life. If voting is socially construed as consenting to join a political community, someone can’t say by fiat that his vote didn’t mean consent any more than he couldn’t say that he shouts racial
epithets because he enjoys the sounds combinations that make up the word. One needs to take responsibility for the predictable consequences of one’s actions when these consequences don’t involve the will of others.\textsuperscript{230} To will the means is to will the end, especially when the means are not strictly necessary in order to live a reasonable life. We can’t get by without living somewhere, but we can get by without voting.

The non-consenting voter nonetheless has one more option available to her. Since her problem is that her act is reasonably being construed in a certain way, she may try to signal in advance that her decision to vote has one meaning and not another. Essentially, she must perform enough symbolic actions of sufficient clarity to signal that it is unlikely that she intended her vote to signal consent. To use an analogy, stretching out while others are signaling "aye" could be reasonably interpreted to count as an "aye." Stretching out and saying "I'm just stretching!" while doing so changes matters.\textsuperscript{231}

Others might be annoyed with this voter for blaspheming what they take to be an important symbolic ritual, but this objection is not of the same kind, nor is it as strong, as the objection that she reneged on her consent. As long as there are sufficiently good reasons for this non-consenting voter to vote in spite of the “blasphemous” character of her action, it will be difficult for others to say that her use

\textsuperscript{230} In other words, there is at least a prima facie difference between Joe’s causing someone offense by screaming racial epithets and “causing” someone to mug him by walking down a tough neighborhood while counting a stack of money.

\textsuperscript{231} The difference between this example and someone’s saying “I’m not racist, I like the sound of racist words” before screaming epithets lies in the plausibility of the statement. We are just more likely to believe the former. If someone insisted that her racist language really wasn’t racist, we would say that either the harm she is causing is enough for her, as a member of the community, to shut up anyway, or that she is so incompetent that she cannot be blamed. In the latter case, group members may reasonably ask whether or not she is “really” an equal member of the community.
of the vote must bind her for reasons pertaining to consent.

This “alternative signaling” can be done individually or institutionally. Individually signaling one’s refusal to allow one’s vote to count as consent is possible but can be fraught with confusion. Except in the most extreme cases, it is rarely clear whether or not someone has done enough to demonstrate that their vote should not be understood as an act of consent. This problem can be overcome if there is a recognized political or social group with which the non-consenting voter identifies and with which she can publicly affiliate. She can, in other words, align herself with an institution that does the signaling for her.

Someone sympathetic to non-consenting voting may object based on their worry that adherence to any official or even recognized political group, including those that advocate separation or revolution, legitimizes the political system in which the group takes part by its mere existence. Yet this need not be the case. We can imagine without difficulty a party dedicated to destroying or radically altering the political and social landscape by any means necessary. It is hard to see how joining such a group would be legitimating the system as a whole if it was well known that it also undertook many delegitimizing actions. Incidentally, if a regime were to ban certain types of parties – including those that would be a home to the non-consenting voter – then the responsibility incurred by the voter to advertise her beliefs diminishes accordingly. This responsibility does not necessarily dissipate altogether, but to the extent that she does not live in a free and open regime, it is attenuated because acting on this responsibility would constitute an unreasonable burden. Thus her lack of
opportunity to express her beliefs via an established party does not mean she must perform supererogatory actions to express her beliefs on an individual level.

Finally, the non-consenting voter's burden increases to the extent that she lives in a free and open society. This is so for at least three reasons. First, she cannot be plausibly expected to announce her intentions if doing so means incurring severe penalties. Her doing so under these conditions would make her heroic, but we since we do not expect or demand heroism, her act would be supererogatory. To the extent that her society is free and open, she does not have this excuse. Second, it is very likely that in such a society she would have many possible avenues by which she can express her beliefs and thus signal a genuine meaning of her vote. Third, to the extent that a society is free and open, people will have good common-sense reasons to believe that normal adults consent to the regime as a general rule. When the political and social climates seem just and conducive to allowing the citizenry to flourish, dissent seems surprising, and so it makes sense to assume that people are consenting unless they forcefully announce that they are not doing so. This belief increases the likelihood that voting will be understood as an act of consent. Dissent is still possible in such cases, but the burden of explanation shifts to the person who, seemingly against all logic as far as the other members of the society are concerned, is dissenting.

IV. Conclusion

Although voting does not necessarily entail consent, there are some cases where it does so. Consent to do what, to whom, and under what conditions? These are
questions that cannot be answered a priori. To the extent that voting is socially understood to constitute consent, then like it or not, it does so. Voting is not necessarily attached to consent, but an individual’s belief that voting somehow does have something to do with consent is neither random nor irrational.

When one votes, one must face the fact that one is creating reasonable expectations. Perhaps this involves the expectation to perform specific actions, or perhaps this "merely" involves a recognition that voting helps constitute or endorse a meaningful shared identity. In either case, earnest consenting voters would be right to feel wronged when discovering that a particular person voted and later announced that the vote should not count as consent. Philosophical anarchists who vote must, therefore, be mindful of what they are signaling when they go to the ballot box.

It should be possible for this non-consenting voter to be able to partake in the other meanings of voting, and this can be done by signaling what her vote is intended to mean. To the extent that she lives in a free, open, and just society, this will be an onerous task if she does so as an individual. She can facilitate the communication, however, by aligning with a party or recognized group which is understood to believe that voting, or at least the type of voting in question, does not constitute consenting.

Even if all this is correct, this chapter's scope is nonetheless limited, as it does not address the glaring omission is the status of non-voters. This is outside the scope of the paper, but I will make two closing remarks on the matter.

First, one strategy on the side of those who argue against philosophical anarchism may be to argue that people have a duty to vote so long as the elections are
free and fair. Singer, in particular, takes this line in his Democracy and Disobedience. The danger here for consent theory is that if one has an obligation to vote on a ground other than consent, then that prior ground (and not consent) would seem to be doing the justificatory work. Political obligation would exist, but not for reasons tied to consent. Second, perhaps it is the case that non-voters should have a different status with respect to their duties toward the state. One intriguing line of thought is that voters may have a duty to obey laws because they have consented to do so (or have through their voting demonstrated that they are part of a collective with a meaningful shared identity and so are bound to follow the decisions of the representatives of that collective), whereas non-voters obey either for other reasons or because they are justifiably coerced into doing so. This latter possibility is especially interesting and should be investigated further. Perhaps the coercion of non-voters would indeed be justified, but it would remain coercion nonetheless. If this speculative line of thought were correct, it would amount to saying that the fundamental connection between citizens may not be as undifferentiated as many democratic theorists hope.

---

Chapter 5: Natural Duty, Dignity, and Respect

Although this chapter describes another theory of legitimacy – natural duty theory – it differs significantly from the previous ones in two ways. First, I do not spend much time trying to unpack it; instead, I mostly just note the philosophical anarchist response to natural duty theories and offer brief commentary. Second, the theory does not “die” with this chapter. While I discussed other theories of obligation with an eye towards moving beyond them, natural duty will be at the root of my subsequent discussion of deliberation, etiquette, and apology. Consequently, this chapter serves as a bridge between the first and second parts of my dissertation. I make this move because I believe that natural duty theories based on deliberation offer the strongest response to philosophical anarchists.\(^\text{233}\)

This chapter is organized as follows. In the first part, I describe a common strand of natural duty theory and explain the philosophical anarchist objections to it. In doing so, I emphasize the importance of dignity in natural duty theories of obligation. In the next part, I articulate one possible, and I believe plausible, relationship between dignity and respect. In the third part, I argue that this type of theory of obligation should, given the right conditions, be agreeable to philosophical anarchists, and that this kind of theory of obligation entails a duty to deliberate with one's possible fellow-citizens.

\(^{233}\) Within the framework of this dissertation, that is. My non-theorist friends have another strong response: "Who cares?" I do not say this flippantly. Philosophical anarchists need to do a better job of explaining why anyone who does not have an intellectual fetish for these questions should engage them. I discuss this further at the very end of the dissertation.
I. Existing Natural Duty Theories and response

I.1 Rawls’s natural duty theory

At their core, natural duty theories of political obligation hold that we have a duty to uphold just institutions. The most prominent articulation of this type of theory – at least in the literature that engages philosophical anarchists – was stated by John Rawls in *A Theory of Justice*. According to Rawls, the most important natural duties pertaining to justice consists of the following: “first, we are to comply with and to do our share in just institutions when they exist and apply to us; and second, we are to assist in the establishment of just arrangements when they do not exist, at least when this can be done with little cost to ourselves.”

The first thing to note is that the duty to obey just institutions is a second-order duty because it only exists to help us discharge more fundamental natural moral duties. In other words, we have natural moral obligations towards others, and if discharging these obligations entails using the state (or other institutions) as a means for doing so, then we must be willing to establish or uphold these necessary institutions as long as doing so does not run afoul our other obligations. Being second-order duties, we must have a substantive account of our moral duties that will give content to the "just" in "just arrangements." Rawls proposes that this substance consist of "justice as fairness," but we need not follow him that far. The second-order duty to obey and establish just institutions (whatever that means) will be enough to cause

---

234 Horton, *Political Obligation*, 103.
236 Ibid., 335.
problems for philosophical anarchists.

These problems mainly come from two facets inherent in natural duties. First, they are non-voluntaristic. We cannot escape these duties or claim that they do not apply to us simply because we have not consented to them. What's more, they are conceptually prepolitical. As Rawls says, “A further feature of natural duties is that they hold between persons irrespective of their institutional relationships; they obtain between all as equal moral persons.” I understand this to mean that the duties exist conceptually prior to the establishment of institutional arrangements and that these arrangements cannot eliminate these duties. I should note here that for Rawls, as with for Kant, this equal moral personhood is an endowment that cannot be lost. If we accept this, then we can see that natural duties of justice provide a moral floor for how we must treat others.

It is here that we can see its appeal for archists, who can demand that philosophical anarchists show that their prepolitical values and duties can be discharged just as well in the absence of a just state. In contrast to voluntaristic accounts of obligation (which philosophical anarchists can say do not apply to those who never consented) or accounts that try to convince philosophical anarchists that they have first-order duties that they did not know they had (such as gratitude theory), proponents of natural duty theories can grant philosophical anarchists their own moral framework and simply ask them how that framework can best be implemented.

At first glance, philosophical anarchists seem to be trapped. If they say that

---

237 Ibid., 114, 334.
238 Ibid., 115.
their duties can be discharged only within the context of their living in existing just states, anarchists can reply that the existence of these just states only came about because most people hold (or at least treat) the state as legitimate. And here philosophical anarchists cannot retreat to the types of arguments that worked in theories of fair play because natural duties – duties which philosophical anarchists endorse – are stripped of all voluntaristic elements.

Recall that fair play theories of obligation such as Klosko’s view the formation or maintenance of the state as a collective project and demands that everyone “pay up” for accepting or receiving benefits. From that perspective, philosophical anarchists can say that they never wanted those benefits or, if they did, that they did not agree to receive them. Natural duty theories, however, constrain philosophical anarchists to acknowledge a duty to participate in the construction of the state in the first place. It's no longer a matter of “we built a thing, now you need to pay up” but “You owe it to us to build this thing, unless you can think of a better, non-hypocritical way of discharging the first-order duties you say you have.”

In essence, natural duty accounts seek to shift the debate away from the fine distinctions made by Wolff and Simmons and back to the kind of ultimatum expressed by Jeffrey Reiman in *In Defense of Political Philosophy*. Frustrated by what he took to be Wolff’s vacuous “anarchism”, Reiman argues that the central question should be whether or not anarchism – actual political anarchism – would better bring about anarchists’ professed values than would a state.239 In other words, he is saying

---

239 Reiman, *Political Philosophy*, 18.
something like “You get to decide what’s important to you and what you think your duties are. Once you’ve done that, do the math and see if what’s important to you can best be instantiated with or without a state. If it’s the former, you’re an archist. If not, you’re a political anarchist. There is no in-between.” This line of argument appears to provide archists with a powerful response to Wolff and Simmons. If proponents of this view are right, then no one can abstain from taking a position on whether or not the existence of the state helps or hurts one’s ability to meet one’s duties.

I.2 Philosophical anarchist response

Although natural duties pose an important challenge, they cannot easily explain why a person should recognize a particular, specific authority as legitimate. Archists do not only say that we have natural duties to states; they claim that we have a special relationship with "our" state. Yet natural duty theory needs more work in order to convince skeptics of the existence of this purported special relationship.

Unsurprisingly, it is A.J. Simmons who brought this point to light. The idea is that natural duties generate duties to human beings due to the latter's status as human beings. They are, as I said earlier, prepolitical. Yet Simmons points out that this inherent cosmopolitanism implies that these duties cannot easily discriminate according to geographical boundaries or tie us to the government that purports to have legitimate authority over us.240 Mere birth in a territory seems like an odd reason to believe that one has duties to the government of that territory.241 There are, for

---

240 Simmons, Moral Principles, 104.
241 Ibid., 148-49. The example that follows is based on his example.
example, many private organizations which (1) are just, (2) apply to me in some sense (for instance, they claim to represent political theorists), and (3) want my support. Were such an organization to suddenly come into existence, or to move to the US from Europe, it would be strange to say that I suddenly find myself with new obligations to that specific entity, even if that entity is a force for good and could help me discharge other duties.

Philosophical anarchists like Simmons have no problem with the idea of natural duties to just institutions. Discussing Rawls’ definition, Simmons says: “Now, I have no quarrel with the second clause of Rawls’s duty of justice, which specifies that we are to assist to some extent in the establishment of just institutions. My difficulties are all with the first, and more central, requirement that we do our share in and comply with just institutions which apply to us.” We can try to save natural duty theory by dropping the "application" clause, but then even philosophical anarchists who live in just states can try to claim that they do not have specific obligations to support the state in which they reside. I will describe two possible ways they have of doing this. The first is intriguing but probably wrong. The second seems to be the more compelling alternative.

1.2.a Supporting non-states

First, a philosophical anarchist may believe that she could best discharge her natural duties by supporting institutions other than states. She might argue that (1) in

---

242 Ibid., 147.
keeping with the latter portion of the 2nd clause of Rawls’ definition, her duties are limited and demand a finite amount of effort (expressed in terms of time, money, or both) on her part, and (2) that she can fully discharge these finite duties by supporting an organization like the Red Cross. She might claim that those who say that she must also support “her” state are asking too much. She may ask why she should allocate her resources away from a just non-state organization and towards the state when the non-state organization seems to be doing a better job at bringing about desirable outcomes. The state might be good, but is it as good as a great non-governmental organization?

This position is interesting but is open to at least two rebuttals. On the one hand, if the philosophical anarchist consistently says that she supports these non-state organizations because they always or almost always do a better job at helping her and others discharge their primary duties than do states, then she is, in fact, a political anarchist, not a philosophical one. Her critic should count this as a victory, even if he must still do the job of arguing against political anarchism.

If on the other hand the philosophical anarchist says that she can provisionally discharge her duties by supporting the Red Cross because doing so is most effective and/or palatable in the current context – one that includes the existence of well-established states – then she can be accused of unfairly free-riding off others' support of the state. After all, as Christopher Wellman points out, this support is what would give the philosophical anarchist the background conditions to allow her pick and

---

243 I assume that the Red Cross is an unobjectionable organization to the reader. If I am wrong, feel free to substitute “Red Cross” with a preferable NGO.
choose how to discharge her duties, and the luxury of getting to choose seems unfair even if she is giving more than is required of her.\textsuperscript{244} Others might rightly say that they wished they had the opportunity to pick and choose where to offer their support and that there is no reason why the philosophical anarchist should have the privilege of making that choice for herself. The crux of the argument is that the philosophical anarchist needs the state to exist in order to be able to support her chosen institutions, and this involves free-riding off the efforts of others.

This response is quite convincing, although even here the philosophical anarchist can try to retort that she is not adopting a privileged position because the question concerns the just and not the good. Her argument might go as follows. Had the philosophical anarchist supported the Red Cross because it represented her conception of the good life, she would be guilty of free-riding on the existing state system. But she might instead support the Red Cross because she sincerely judges that doing so would help advance a conception of justice that any moral person would endorse.\textsuperscript{245} Whether or not this is correct, it is not free-riding.

She may reason that since most people happen to support the state for whatever reason, her doing so would be redundant since the state could in principle absorb a certain number of people like her without becoming unstable or unable to perform its functions. Had she spent her time or money supporting the state, then justice would require that someone else take her place and support the Red Cross or its equivalent.

\textsuperscript{244}Wellman and Simmons, \textit{Duty to Obey}, 42.
\textsuperscript{245}Foreshadowing the next main section of this paper, she might say that it helps protect the universal rights that stem from human dignity.
In other words, her support for the Red Cross may appear to be an unfair differentiation, but the fact that she is supporting it instead of the state is irrelevant. Money is fungible, so it wouldn’t matter which bucket my dollar goes into as long as it goes into a bucket that must be filled as a matter of justice. To those who propose that she support both institutions, the philosophical anarchist could reply that natural duties of justice are finite and that there is only so much we can demand of her and others.

Of course, this is not the final word on the matter, and I do not propose a definitive answer here. A response to the defense offered above is that the philosophical anarchist’s presumption of the right to use her discretion is problematic because (1) it necessarily draws on controversial opinions – is the Red Cross better than Doctors Without Borders? Can time be substituted for money? – and (2) it prevents others from exercising their discretion without committing injustices, i.e., withholding their now-necessary support of the state.\footnote{Wellman seems to think that “each of us has good reason to want to be the author of our own lives” for reasons that go beyond the fact that we would presumably use that discretion to optimize our choices. The authorship itself seems to be a desired good. Wellman and Simmons, \textit{Duty to Obey}, 41.}

Philosophical anarchists might respond to this by digging in their heels and insisting that one of our duties – the primary duty, to many – is to make our own determinations and to act accordingly.\footnote{Again, this is Wolff’s position. See Wolff, \textit{Defense of Anarchism}, 12-14.} After all, they seem to think that doing so is essential to autonomy. Yet these same philosophical anarchists may also recognize that making our own determinations involves acknowledging our fallibility and that we need political institutions to collectively decide which other institutions to support.
in order to discharge our first-order duties. This would cause the philosophical anarchists to at least acknowledge that they should try to set up these institutions. This is essentially the line that I take in this chapter and the remainder of the dissertation. Philosophical anarchists, in fact, have a duty to at least try to engage others in order to discharge their natural duties more effectively, and this engagement may result in binding agreements. This engagement may fail, and there will be other reasons that undergird the duty to engage, but the upshot is that philosophical anarchists have a choice: jettison morality, keep morality and become political anarchists, or acknowledge a duty to attempt to deliberate with others.

1.2.b Supporting other states

Finally, what about the philosophical anarchist who accepts all this but says that none of the preceding discussion implies that she must support her state? Why should she pay taxes to "her" government instead of a neighboring country's? Even if she acknowledges the need for government and that government involves states, she may reason that the money should go to the neediest states. After all, her tax dollars would go much further in Malawi (assuming they were spent more or less efficiently) than in the US. Or perhaps she is impressed by policies like No Child Left Behind and decides to fund the best performing states and withhold money from the rest. As an American, she might think that the US is merely basically just but that Japan really gets justice right. Why keep dumping money into a barely acceptable system when she

---

248 Simmons made this point in his essay, The Obligations of Citizens. See Simmons, “Justice and Legitimacy,” 47.
could vote with her dollars and support a better one?

The best answer to this line of questioning begins with the observation that globalization notwithstanding, our actions mainly affect those around us. To the extent that this is true and that we live in conditions similar to what Rawls calls "moderate scarcity," living as we please runs the risk of stepping on others' toes. Wellman expresses this worry and uses it to justify the need to have a common set of rules, but one need not accept this conclusion to accept that proximity leads to friction and an increased chance of doing injustice to others. Basic consideration of others as human beings – the same consideration that informs our natural duties towards them – impels us to at least try to communicate in order to achieve basic coordination, even if this coordination amounts to an attempt to fully stay out of each others' way as we independently pursue non-vicious life projects. This, I will emphasize at the end of this chapter and in the following chapter, is what motivates a duty to at least try to deliberate.

Summing up: modern philosophical anarchists have accepted the existence of natural duties. Following Kant or Rawls, we have reason to think that discharging these duties may, in turn, generate duties to establish or support just institutions that allow us to better meet our primary duties. Yet natural duty theories have trouble

---

249 I accept that this is subject to debate. I should note that I have no immediate reason to reject the contention that those non-citizens that a country's policies affect should have some say in the political decisions of that country, even if they live abroad.

250 Rawls describes moderate scarcity as follows: "Natural and other resources are not so abundant that schemes of cooperation become superfluous, nor are conditions so harsh that fruitful ventures must inevitably break down." Rawls, Theory of Justice, 127.

251 Wellman and Simmons, Duty to Obey, 38.

252 In some cases, we may have a duty to try to thwart the vicious designs of others.
morally binding citizens to their state. Some might argue that natural duties are best discharged by supporting non-state institutions, and others might argue that the duties are best discharged by supporting the neediest or most deserving states.

Philosophical anarchists seem to have the duty to at least try to engage with others, both to see if coordination would allow them to better discharge their duties as well to try to avoid inadvertently harming those within their vicinity. In the remainder of this chapter, I will try to explain in greater detail the source of this duty while specifying some of its limits. Ultimately, I want to answer two questions. The first is, "What must philosophical anarchists do?" The second is "Under what conditions can philosophical anarchists say that their duty to engage has been satisfied, allowing them to retreat to their default position that there is no prima facie duty to obey the law?"

II. Dignity and Respect

The main appeal of natural duty theory is the intuition that we somehow owe others something by virtue of their mere personhood. And as Simmons has argued, its main weakness in the context of this discussion stems from its difficulty in explaining why we are bound to particular, specific governments. To say that I owe others something independently of any explicit agreement seems reasonable; saying that I must respect the authority of a particular state – one which I might find significantly less just than another organization – seems to be a bridge too far philosophical anarchists.

In the remainder of this chapter, I propose one way of trying to save the baby
of natural duty while dumping the bathwater. My essential argument is that natural
duty theories are (or should be) grounded on a conception of human dignity and that
this dignity, while universal, gives rise to two types of duties. Some duties are
universal and must be discharged in order to treat others with the respect that flows
from our recognition of their having dignity. This will ground the duty to try to
deliberate with others. Other duties, while also stemming from our status as beings
with dignity, demand differential treatment and will prove to be the source of
deliberation's fragility. Put another way: the existence of natural duties both supports
and undermines the claim that the state has authority over us.

In making this argument, I have no pretensions of either revolutionizing the
discourse on dignity (or respect) or coming close to surveying the literature on these
topics. My goal is limited: to propose an account that is (1) plausible, (2) consistent,
and (3) in line with the frameworks of the major modern philosophical anarchists,
notably Simmons and Wolff. To the extent that I succeed, I will show that
philosophical anarchists have a duty to deliberate, a duty which may in some cases
constrain them to accept the legitimacy of the state. Yet I will argue, mainly by way of
the final two chapters of this dissertation, that many philosophical anarchists may
avoid becoming archists while still discharging their duties. Still, I believe that if they
take their duties seriously, they must engage in interactions which could entail
accepting the legitimacy of the state, whether the (now former) philosophical
anarchists like it or not.
II.1 Dignity

The problem with using dignity as the groundwork for any theory of obligation is that it comes very close to being what Walter Gallie calls an "essentially contested concept", that is one that is subject to many different uses by people who often assume they are talking about the same thing.\(^{253}\) I will not try to adjudicate between conceptions of dignity, nor will I spend much time discussing its fine points. The purpose of this section is to put forth a basic idea of dignity which I will connect with the idea of respect in the section II.2.

In pursuing this aim, I will treat dignity and respect as conceptually distinct and will resist attempts to conflate them. This is because I will ultimately argue that the respect, but not dignity, that we show others is in some cases (but not others) based on our assessment of how they act. So while dignity will be universal and stable, this will not always be the case with respect. As a result, I avoid terms such as Nancy Sherman's "dignitary respect."\(^{254}\) She presents this type of respect in contrast to the "deference respect" that members of the military owe their superiors based on rank. Sherman argues that "dignitary respect," on the other hand, is the type of a more "democratic form" of respect motivated by the belief that "... such deference rituals, within the military and elsewhere, ought to be undergirded by modes of mutual respect due all persons, independent of age or status."\(^{255}\) It is, in other words, the kind of unearned respect that we are all entitled to.

---

\(^{253}\) Gallie, “Essentially Contested Concepts.”
\(^{254}\) Sherman, *Stoic Warriors*, 51.
\(^{255}\) Ibid. Sherman’s use of the term “dignitary” is unfortunate, given that it is a term usually used to denote individuals of high rank.
Sherman’s point – that some forms of respect like dignitary respect should emerge from a conception of universal human worth – is ultimately valuable but should only be entertained after we make a clear demarcation between the two terms. To see this, we need only think of the ways that our use “respect” and “dignity” causes us to slip across concepts. To say that a person “always acts in a dignified manner” implies a contrast between that person and the rest of the cursing, anxious, and at times slovenly crowd. It is the opposite of the egalitarian dignity that we are owed by virtue of our personhood. Someone might retort that she, unlike the rest of us, is acting in a manner befitting a person with dignity, and that since this manner of acting is available to all, the seemingly divisive use of “dignified” in fact points to a common standard by which we should be judged. Yet this seems for two reasons.

First, what counts as “dignified” is open to contestation, as one person’s dignified manner is another’s snobbery. Someone who is dignified will engender differing levels of respect from others based on their assessment of that person's demeanor. Second, this understanding of "dignity" runs counter to the prevailing conception and use of the word as something that all of us, or at least most of us, possess. The point is that while there is surely a connection between dignity, dignified behavior, and respect, conflating or combining the terms only opens a conceptual

---

256 Watts, Politeness. I will have more to say on politeness in Chapter 7 of this dissertation. For now, the important point is that many people would interpret politeness – including some forms of dignified behavior – as a mechanism of oppression, with those in the upper classes both (1) dictating what counts as dignified, and (2) promulgating the idea that such behavior is available to all who have aspiration of class ascension. The trick, however, is that the rules of polite behavior, as well as dignified behavior, changes as the assessors of dignified behavior move to stymie the hordes of parvenus who are trying to take their place in their class stratum.
Pandora's box. It is better to conceive of the connection in a step-wise fashion: first, we define dignity, then we define respect, and consider the relationship between the two.

The conception of dignity that I use here is simple and has a wide reach. I follow Kant in saying that beings with dignity have "an unconditional, incomparable value", and that, as Michael Rosen says, "we could simply use that latter phrase to replace most of Kant's uses of the term 'dignity'." It is true that Kant intended this term to apply to (1) all human beings and (2) only to human beings, and his influence is such that even if (2) has recently come under mild attack by animal rights proponents, few people outside of a subset of professional philosophers would deny (1). Even if we grant that some non-humans have dignity, what is important – and what drives our reluctance to “grant” dignity too easily to other species – is that beings with dignity necessarily have equal dignity, as the idea of dignity as “incomparable value” makes it inherently egalitarian.

One need not accept Kantian metaphysics to accept this definition. In fact, I am counting on this fact. The advantage of having a broad working definition of dignity is that it could help support an argument that speaks to a relatively wide range of philosophical anarchists, including Wolff – an odd kind of Kantian – and Simmons, an equally odd Lockean. What is important for my purposes is that (1) we have dignity

---

257 Rosen, Dignity, 21. Kant’s most statement here is: “In the kingdom of ends everything has either a price or a dignity. Whatever has a price can be replaced by something else as its equivalent; on the other hand, whatever is above all price, and therefore admits of no equivalent, has a dignity.” Kant, Metaphysics of Morals, 40.
258 For example, Regan, “Animal Rights.”
259 Rosen, Dignity, 24.
by virtue of our being persons, and (2) that dignity is binary – either one has it or one
doesn't, and (3) that our status as beings with dignity means that we are owed certain
types of treatment by those who can recognize this dignity.

Dignity as I understand it provides the foundation of at least some (and
possibly all) moral duties. This means that I resist some of George Kateb’s account of
dignity. Kateb claims that dignity is an existential value, not a moral one:

All existential values have a conceptual independence from
instrumental practicality and most important from morality, despite the
fact that freedom and equality, the core of human rights, are often
defended as practically or morally necessary or useful. [...] I mean that
for many people, and rightly, morality has to do solely or principally
with human suffering; but human dignity in its concern with status and
stature has to do with the proper recognition of the identity of every
human being and the identity of the human species.²⁶⁰

I disagree with Kateb's blanket assertion that morality "has to do solely or principally
with human suffering," but agree with the thrust of the latter part of the quote. Dignity
is concerned with our stature as persons and the recognition of that stature. It involves
what Michael Rosen calls "respect-as-observance", or the recognition of others as
having dignity.²⁶¹ This stature cannot be bestowed or taken from us.²⁶² Finally, not
only are we a certain way, but our being that way compels others to see us as having a
certain stature. Again, this is not the kind of stature that increases or decreases as we

²⁶⁰ Kateb, Human Dignity, 12-13. In order to avoid misrepresenting Kateb, I should point out that it
would be a mistake to take this passage to mean that he is simply a utilitarian. For a more complete
account of Kateb’s view, see Human Dignity, 82-92.
²⁶¹ Rosen, Dignity, 57.
²⁶² Clarence Thomas expressed this thought, while implicitly (and probably inadvertently)
highlighting the tension between the different uses of dignity, when he wrote in his dissent to the
Obergefell v. Hodges case, when he said: “Slaves did not lose their dignity (any more than they
lost their humanity) because the government allowed them to be enslaved.” Obergefell v. Hodges.
This naturally drew heavy criticism from those who thought that such rhetoric excused the vices of
slavery.
perform morally or socially laudable (or shameful) acts. That kind of status is tied to a
type of respect which must be *demonstrated* – similar to what Rosen calls “respect-as-
respectfulness” in which we *show* others respect – to which I turn in the next
subsection.²⁶³

Before doing so, I want to linger on Kateb’s concept of dignity a bit longer in
order to draw out some important ideas that I have not yet made explicit and that will
be useful later. Kateb’s most succinct description of dignity is as follows: “All
individuals are equal; no other species is equal to humanity. These are the two basic
propositions that make up the concept of human dignity.”²⁶⁴ To take the latter
proposition first, we are bestowed dignity in part by virtue of our belonging to a
species which, Kateb believes, has made a partial break with nature and can use “its
partial nonnaturalness in the service of nature.”²⁶⁵ In Kateb’s conception, one desirable
way that we as a species can express our distinctive traits is by acting as nature’s
steward, that is, to use our mental and aesthetic attributes to treat nature in a way as it
should be treated if it had the ability to know how we are caring for it.²⁶⁶ We are to do
this not merely to “exult in uniquely human powers,” but to admire it as a form of
gratitude for the “wonder” we feel when we contemplate existence.²⁶⁷ Human beings,
apart from all other species, are able to do this.

More importantly for my purposes, however, is Kateb’s concept of the dignity

²⁶³ Rosen, *Dignity*, 58.
²⁶⁵ Ibid., 115.
²⁶⁶ Ibid., 117.
²⁶⁷ Ibid., 118.
of the individual within the species. His assertion that we have equal dignity is in line with what I have written so far. Yet while I have so far described the egalitarianism of dignity in a way that downplays difference, Kateb’s story emphasizes our distinctiveness. As human beings, we may develop a self-consciousness that gives rise to free agency – once again marking a break with nature – and while some may not develop self-consciousness, those people still have the potential to do so.\textsuperscript{268} After all, Kateb reasons, what distinguishes us from nature and forms part of the basis for our dignity is our inability to, in good faith, describe our identity or that of others in terms of a closed-ended totality.\textsuperscript{269} In the proper conditions, “a particular self-consciousness will develop; the ego will grow in a certain direction. One feels special not because one has what others lack, or has a rank higher than others, but because one has what everyone is entitled to have, just by being human.”\textsuperscript{270} We are thus equal, but this equality within our species is based on our lack of finitude, which effectively prevents us from undertaking honest existential comparisons between us.

Kateb is keen to emphasize our uniqueness both in terms of our potentiality as well as in terms of our perspective. He defines a perspective as follows: “A perspective is what happens to a person without the person knowing it, an unresisted result of background and temperament, until the person notices that other persons look at the world differently, whether in one’s society or in other societies.”\textsuperscript{271} This difference is not superficial – Kateb says that it is “impossible” for a person to know

\begin{flushright}
\textsuperscript{268} Ibid., 162-163. \\
\textsuperscript{269} Ibid., 134. \\
\textsuperscript{270} Ibid., 163. \\
\textsuperscript{271} Ibid., 146. 
\end{flushright}
what it is like to be another person, and that “empathetic interpretation goes only so far, and often not very far.”272 Each of us is uniquely situated to experience our inner life, of our being who we are.273 Yet neither is the difference absolute. Kateb believes that recognition of our shared membership in our species gives us a common framework upon which we can attempt to understand and appreciate our differences.274 In other words, our commonality allows us to highlight, not wash away, what makes us individually unique.

Here someone might ask if all this implies that Kateb is an anarchist, philosophical or otherwise. For one thing, Kateb’s sympathy toward the importance of the unique individual resonates nicely with the overall tone of philosophical anarchists like Wolff and Simmons. What’s more, a standard anarchist critique is that governments suppress individuality and diminish our capacity to develop a robust self-consciousness (in Kateb’s sense), charges that would imply that supporting the state would constitute supporting an institution that violates the dignity of its citizens.275, Kateb further expresses a deep suspicion for the kind of group identification held by many archists; in “Is Patriotism a Mistake?” he writes:

Patriotism (and some other group affiliations) signify that a person, as oneself and in oneself, is not very much. Yet by giving oneself to the group, one does one's share in keeping alive an entity that can be or become, as itself and in itself, something that matters greatly and is worthwhile. Patriotism, more than almost all other group affiliations, is a way of acquiescing in one's inferiority. What Emerson says about popular loyalty to kinds and heroes can be transferred to an impersonal

272 Ibid., 152.
273 Ibid., 130.
274 Ibid., 168.
275 For example, see Goldman, Anarchism and Other Essays, 70-71.
entity like a country.\textsuperscript{276}

Kateb, however, is no anarchist. While he is clearly suspicious of the state, he also endorses its existence. He believes that it is under the auspices of a state that protects individual rights that we are able to develop the type of self-consciousness that we are entitled to as human beings.\textsuperscript{277} And while warns that governments will predictably attempt to overstep their appropriate bounds under the guise of trying to protect society, he also concedes that “by providing security, government makes possible treating other persons morally (and for their own sake).”\textsuperscript{278}

So where does this brief description of Kateb’s thoughts on dignity and the state leave us? Kateb’s framework reminds us that an egalitarian dignity that unites all people does not imply that we will come to the same conclusions or understand others if we merely speak to them long enough – a point that will become especially salient in the next chapter, where I discuss deliberation. Our shared humanity makes us entitled to certain types of behavior from others and gives us a framework to view others as something other than alien, but, as per Kateb, we should not overestimate our capacity to empathize with others or to make them understand our perspective. As human beings we have equal dignity, but this equal dignity does not imply sameness.

One need not be a follower of Kateb for the remainder of my argument to stand. His framework provides one way out of many of connecting universal human dignity with uniqueness and a skepticism toward those who would use the idea of

\textsuperscript{276} Kateb, “Is Patriotism a Mistake?” 16-17.
\textsuperscript{277} Kateb, \textit{Human Dignity}, 163.
\textsuperscript{278} Kateb, “Is Patriotism a Mistake?” 19.
dignity to smuggle in a view of people as homogeneous and thus commensurable. As I will discuss in the following subsection, our status as beings with dignity does confer universal, undifferentiated duties. It also sets the stage, however, for cases where we ought to show others differentiated respect.

II.2 Respect

While dignity is monolithic, universal and egalitarian, with each human being holding an equal (or incomparable) "amount" of it, respect can be divided into subclasses. One type of respect flows from our concept of dignity and concerns the attitude and actions we are constrained to adopt when we consider or interact with someone with dignity. This type of respect, like dignity, is universal. The other type of respect that concerns me is the type of response that we believe people we are owed by virtue of their actions or beliefs. Disrespect of our mere status as individuals is obviously a moral wrong. Disrespect based on an assessment of our actions or social status is trickier; I will argue that it is only wrong in some cases. The upshot will be that universal dignity generates duties that make us come together, while some types of differential respect are also connected with dignity (in an attenuated way) but push us apart.

Before beginning, it may help see what is at stake by considering a series of questions. Why might so many people feel awkward when, while waiting at a bus stop, someone begins to sort through a nearby trash can in search of cans to recycle? Let’s ignore dodges like “I feel unsafe,” although that answer is worthy of follow-up questions in its own right. Questions of safety aside, most of the people that I have
informally polled confirmed that they do feel awkward in these situations.

I am willing to bet that those who feel awkward in these cases would feel a different kind of awkwardness when the scenario occurs while they are visiting a foreign country. When one is abroad, the tension reads like “I am a good person and a cosmopolitan – how can I reconcile that with the fact that I am not doing as much as I could to help people like this?” In other words, “I know that I am the beneficiary of an unfair circumstance.” Yet even if the future bus passenger subscribes to some sort of cosmopolitanism, there is still an element of otherness which makes the interaction easier. She may know that she somehow participates in the perpetuation of global injustice, but the feeling of responsibility is normally attenuated enough to allow feelings of pity to replace most feelings of responsibility.

I would speculate that when we are at home (whatever that means to us), our feelings of discomfort are colored by a feeling of uncertainty. More specifically, it feels disingenuous to adopt the same “caring but separate” attitude as in the case of the tourist. When we walk down our streets, we occasionally accidentally make too-prolonged eye contact with passerby, and these potentially awkward encounters are oftentimes diffused by a head nod or a mumbled “Good morning.” I will have much more to say about these types of encounters in a later chapter, but for now suffice to say that the nod and mumble indicate a basic egalitarian recognition, thus diffusing any question of possible hostility. This option seems more difficult to pull off at the bus stop, though. Not only is the scene stationary – the person waiting for the bus can’t “keep walking” – but the nod is not part of our standard behavioral repertoire.
I suggest that this is because the desire to express a recognition of the other as an equal is belied by the mutual knowledge that the person sorting through the trash is, in spite of the best intentions most, seen as engaging in a shameful act. And unlike the case of the tourist, the person who is waiting for the bus “at home” implicitly recognizes the person looking for cans is part of her political community. Hence the feeling of uncertainty – what does the bus passenger do to express her recognition of the other as an equal without insulting both of them by pretending that she is unaware of the mainstream social significance of picking cans out of the garbage? What does she owe this person, and how does she give him what he is owed? And to what degree is she responsible for the tension felt when someone with equal dignity is performing an act that is deemed shameful?

In sum, the person waiting for the bus may become embroiled in a complicated, painful internal struggle where she recognizes the dignity of the other but doesn’t know how to express it without disrespecting him. Were the question only about mere dignity, she could pity him with minimal complications. But this case is not only about dignity – it’s about the tension stemming from conflicting demands made by our recognition of others as beings with dignity, and the claims they implicitly have as members of our political community. These are the issues that should be kept in the background and that inform the rest of this chapter, and the next three chapters as well.

II.2.a Universal respect
Kant best describes the type of respect that we are owed by virtue of our having dignity. We can again turn to Rosen here: “What is to be recognized as worthy of respect, according to Kant, is not the position that an individual occupies within a particular society […] but the lawgiving function of morality, something that human beings carry inalienably within themselves.” Kant, of course, attributes our dignity as individuals to our autonomy – our ability to be the authors of our own laws – but what matters here is not so much the specific ground of dignity that Kant defends, but rather the fact that the respect we are owed is due to (1) some attribute of persons (that is somehow connected with autonomy) which (2) is discernible by other persons. This discernment is important, as respect in this sense rests on our capacity, indeed, our compulsion, to recognize other persons qua persons. Stephen Darwall’s most recent book, *The Second-Person Standpoint*, is especially helpful here. Not only will it help clarify what I mean by "discernment," but it will also serve as a basis that will ground a duty to deliberate.

In *The Second-Person Standpoint*, Darwall claims that “the second-person standpoint” should serve as the foundation for deontological ethics. He describes this standpoint as “the perspective you and I take up when we make and acknowledge claims on one another’s conduct and will.” This is a particular type of acknowledgment, one that differs from many other types of responses that are connected to moral considerations. To illustrate by way of what is apparently his favorite example, if someone steps on my foot, my saying “Hey! Get off my foot!”

---

279 Ibid., 26.
appeals to something other than the pain I may feel as a result of the person’s action. To be sure, the foot-trampler could be persuaded by the knowledge that I am in pain. He might reason that more pain is morally bad in a generic sense and that his action contributes to this bad state of affairs. When this type of person hears "Get off my foot," they are in effect interpreting my utterance as an epistemic argument. Put another way, "Get off my foot" provides information regarding pain that was otherwise unavailable.  

Darwall points out that this type of reasoning differs from what is going on when I make the claim that the person ought to get off my foot. In making my claim, I may still give reasons – such as the fact that my foot is in pain or that my once-spotless shoes have been smudged – but in doing so, I am doing more than attempting to cajole or point out objective considerations. I am presupposing that I have the right to make a demand. This is the essence of Darwall's second-order reason. It is agent-relative as opposed to the "objective" or "agent-neutral" types of utterances inherent to what Darwall calls the third-person perspective. As such, the second-person standpoint is about personal, not logical, relations. It also differs from first-personal thought, which “lacks an addressing, second-personal aspect.” It is neither pure self-reflection nor pure contemplation of the moral landscape as a whole. It is a reason addressed to a person.  

281 Ibid., 5.  
282 Ibid., 9-10.  
283 Ibid., 10.  
284 That said, the reason need not necessarily be uttered. Fear, awkwardness, shyness, politeness, or any other number of factors may cause me to avoid actually saying "Get off my foot". Still, if I am getting indignant, I will in some way be performing the "Get off my foot" routine inwardly.
Any second-order reason thus makes a number of presuppositions, not the least being that (1) the person giving the reason has the authority to do so and (2) that the person hearing the reason has the capacity to understand and act on it. "A second-personal reason is one whose validity depends on presupposed authority and accountability relations between persons and, therefore, on the possibility of the reason's being addressed person-to-person."\footnote{285 Darwall, \textit{Second-Person Standpoint}, 8.}

What’s more, Darwall believes that dignity is tied to second-order reasons. "… what is called the inviolable value or \textit{dignity} of persons has an irreducibly second-personal element, which includes the authority to demand certain treatment of each other, like not stepping on one another’s feet."\footnote{286 Ibid., 13.}

If Darwall is right, then dignity is not simply about discerning the rights of others but about the demands we have the authority to make. Against an atomistic, pre-relational conception of dignity, Darwall says:

\begin{quote}
Someone might accept the first-order norms that structure the dignity of persons and regulate himself scrupulously by them without accepting anyone’s authority to demand that he do so. He might even accept these as mandatory norms in some suitable sense without accepting anyone’s claim to his compliance. I claim, however, that he would not yet fully acknowledge the dignity of persons or respect persons for their dignity. These involve an irreducibly second-personal dimension.\footnote{287 Ibid., 14.}
\end{quote}

In defending this second-person perspective, Darwall seeks to distance himself from Kant, who he says grounds the categorical imperative a conception of autonomy that necessitates the presumption of “the idea of freedom” when engaging in \textit{any} practical
action. But Darwall believes that it is possible to imagine practical reasoning *without* presuming autonomy. Instead, he grounds autonomy on *deliberation*: “The problem, as I diagnose it, is that it is only deliberation from the second-person standpoint that requires us to assume autonomy of the will.” This deliberation presumes second-person *competence*, which “must consist in something like the capacity to make demands of oneself from a second-person standpoint: in being able to choose to do something only if it is consistent with demands one (or anyone) would make of anyone (hence one would make of oneself) from a standpoint we can share as mutually accountable persons.” The idea of accountability is central here, and the move towards universalism implies an identical set of duties that each person owes everyone – including herself – that stem from that person’s dignity, which in turn stems from their competence as second-person actors.

Here I want to pause and clarify the implications for my argument. If Darwall is right, deliberation is clearly tied to the essence of dignity and the respect we show others that stem from this dignity. It would mean that treating people as people requires addressing them while recognizing and respecting their ability to make autonomous judgments. The recognition can’t be helped; the very act of making a moral claim on another implicitly does the job. One can fail to respect others’ dignity, however, by deliberately *disrespecting* their status as persons, that is, to treat them as though they were not beings capable of making autonomous judgments or did not

---

288 Ibid., 30.
289 Ibid., 32.
290 Ibid., 35.
possess second-person competence.\textsuperscript{291}

If you do not buy Darwall’s argument or the implications I claim stem from it, however, Kant’s theory serves my purposes as long as it is understood to contain a duty to attempt to deliberate with others. In fact, any natural duty theory which contains this particular duty as part of a universal set of basic duties stemming from dignity should do the job. Were the reader to accept the duty to deliberate at the outset, this entire part of the discussion would be irrelevant. I include it in order to make the duty to deliberate more intuitively plausible, not to construct an argument that will satisfy all flavors of Kantianism. That said, even if Darwall’s argument is unpersuasive to those who prefer these other flavors, other justifications may satisfy them that such a right exists.

In concluding this subsection, I would like to make three additional points. First, this duty of respect includes self-respect. We may be no more important than others qua humans, and we may owe them a great deal, but we also have significant duties towards ourselves. This means that there will be a limit to the amount of deference we can show others without fundamentally disrespecting ourselves. Self-respect is, of course, compatible with Darwall's second-person standpoint; much of the moral life of each person is that person's internal dialogue with himself.

Second, the respect we owe ourselves and others is based on some capacity to conceive of and act according to a moral schema. In other words, if you take seriously my demand that you get off my foot, you do so because you recognize me as a person

\textsuperscript{291} It is thus the opposite of Baugh’s portrayal of a Wolffian agent, which I describe in the final chapter of this dissertation.
who can connect that demand to a set of non-arbitrary assertions of how we ought to treat each other. Our status as moral beings thus surely entails an obligation to (1) determine what does and does not fall under the scope of the moral sphere, (2) to determine the best course of action, and (3) to act according to our best determinations. This does not mean that any given individual will be correct, or even that there necessarily is an objectively correct moral framework. It simply means that everyone has a duty to try to figure out and do the right thing (when there is a right thing to do, that is) and that any given individual will attempt, as an earnest actor who has the capacity to act morally and who takes her moral duties seriously, to take the course of action that she believes to be best.

Third, and most obviously, we have universal duties towards others. The nature of these duties will surely be in dispute. Kant, for instance, famously divided duties into *perfect* and *imperfect* duties. Perfect duties are usually negative duties can never be violated – injunctions against lying or murder, for instance. Imperfect duties are usually positive duties and allow individuals a certain degree of discretion in determining how to discharge them. Kant believed that we had a duty to assist others, that is, a duty of charity, but did not believe that this entailed a duty to help

---

292 I am not masochistic enough to enter *that* debate.

293 In fact, if there is some objectively correct path and an individual mistakenly comes to believe that she ought to take another one, it is hard to see how she can avoid moral failure. Taking the mistaken path causes an objective bad, but performing the correct action, possibly after some browbeating by her neighbors, risks having her fail to abide by her duty to take the action that she believes to be correct. Moves such as introducing maxims like "do what others suggest when a vast majority of them take the same position" don't help since a person who is trying to be moral would only adopt that maxim if she believed that following it (in opposition to her gut) is the right thing to do. To use a mundane example: I might come to accept that some behavior of mine is hurtful even if it seems innocuous to me if enough people tell me that I am being a jerk.

every person that we encounter who needs help. Wolff, as I show in Chapter 9, disagrees that there exist any positive duties and that our duties towards others are strictly negative. In spite of these differences, however, anyone who endorses the idea of human dignity acknowledges that the presence of others immediately constrains what we are allowed to do. The presence of others, in other words, presents us with a problem. To the extent that the set of moral duties that we have derived prior to our encounters with others fails to regulate our behavior in a way that mitigates conflict, we will have a universal duty to try to resolve these frictions in a way that demonstrates respect for ourselves and others. Part of the moral framework that we have a duty to construct must include mechanisms designed to handle the rough spots caused by our interactions with others.

II.2.b Differential respect detached from dignity

The second type of respect concerns the valuation we believe we are owed relative to others and can be thought of as "differential respect," which can be usefully subdivided into two further categories. One version of this is unconnected to or contrary to our sense of dignity. The other stems from our sense of dignity but, unlike the type of respect described in the previous section, does not give rise to universal duties. This subsection discusses the former, and the next subsection discusses the latter.

The first type of differential respect is the kind of respect demanded by celebrities who angrily ask, “Do you know who I am?” when they don’t get their way.
By asking this question, the celebrity is implicitly saying that he is someone more important than others and should receive special treatment. At best, the celebrity may simply be threatening the other person, e.g., "Do what I want or you'll be sorry," while recognizing their fundamental equality. At worst, the celebrity may be implying that he has more intrinsic worth than the recipient of his bile. And somewhere in between, he may feel that although they are both persons, the celebrity should nonetheless be accorded special deference due to his stature.

A second, even more obvious example is the respect that a slave owner demands from her slaves. Under many conceptions of slavery, the slave owner simply denies that the slave has dignity. In others, she may acknowledge that the slave has dignity but that the slaveowner should be accorded additional rights and deference as a result of her social position.

These types of arguments surely do not generate prima facie duties. Those who deny the dignity of the other individual are wrong by hypothesis, at least as far as this dissertation is concerned, so I will pass over them. Those arguments that claim equal dignity between persons while calling for differential respect are more

295 The person might adopt the attitude that “anyone” could or should take advantage of their dominant position. This need not involve feeling entitled; it could be the result of seeing life as a kind of game where anyone with good sense would exploit their position. As International Chess Master Jeremy Silman puts it: “If the board calls for a kingside attack, then you must attack.” Silman, The Amateur’s Mind, 297.

296 They may, however, produce circumstantial duties.

297 I acknowledge that many have argued against the existence of dignity. Arthur Schopenhauer perhaps put it best: “That expression, dignity of man, once uttered by Kant, afterward became the shibboleth of all the perplexed and empty-headed moralists who concealed behind that imposing expression their lack of any real basis of morals, or, at any rate, of one that had any meaning.” Rosen, Dignity, 1. That said, I take for granted that the reader is at least willing to entertain the possibility that dignity exists. I admit that my entire account would be severely weakened – possibly defeated – in the face of a strong argument against the existence of dignity. So much the better for committed philosophical anarchists.
interesting. Sometimes these claims have nothing to do with dignity and are seen as circumstantially appropriate. For instance, we may abhor a celebrity who cuts in front of us in line using the "do you know who I am?" routine but recognize the appropriateness of the demand as she tries to enter an award ceremony where we agree she "belongs." Or, to use another example, we might say that a superstar political theorist deserves (more) respect (than her peers). Those who failure to appreciate her status could be seen as lacking in of sophistication at best and being rude at worst. We might even derisively say that the hotshot young political theory hire is acting “entitled” when he acts as though he should get as much consideration as the acknowledged superstar at the conference.

These examples share important traits. First, the claims to differential respect are circumstantial. Both the celebrity and political theorist superstar believe they deserve deferential treatment within a narrow social sphere. For their claims to be plausible, they must avoid implying that their elevated status extends to their life as a whole, or, worse yet, to their personhood. The superstar theorist may expect and even feel entitled to special treatment at a conference, but she hopefully does not feel that this special treatment should extend to most of her day-to-day interactions.

Second, these demands for differential respect are not tied to dignity in a clear way. This is in large part why they are circumstantial. The actor deserves respect in contexts where his professional status is relevant, and the same is true of the political theorist. The matter becomes muddied, however, because the demands for differential respect are based on traits or circumstances which, if not obviously tied to dignity, can
still be admirable. For example, the celebrity might claim that he ought to be admired and treated differently because his success is a reflection of the desirable traits that he possesses, such as a willingness to work hard, be personable, and look good.

These traits are clearly not created equal, at least from a moral perspective. Many people would say that hard work and geniality are morally relevant, while few would say the same of getting a high rating at hotornot.com. Those who acknowledge the equal dignity of persons and who want differential respect and who want the respect to transcend the particular social context where their unique traits are relevant must either find a reason that does not tie the respect to dignity or must think of some way of saying that equal dignity should nevertheless lead to unequal treatment in some cases.

The first strategy, which does not tie respect to dignity, represents respect as something divorced from most conceptions of morality, and certainly from conceptions tied to natural duty. A defender of this perspective might say that dignity gives us a baseline for how we ought to treat ourselves and others, but that other considerations are important in that they give our lives meaning. On this view, the political theory superstar might admit (if pressed) that there's nothing morally laudable about being a professionally successful political theorist while still maintaining that she still thinks that others should look up to her because political theory is connected with desirable, but not morally relevant, traits. In doing so, she is likely to either try to popularize political theory, to increase the prestige of the discipline, or to claim that success at political theory implies that she holds certain desirable but ultimately non-
moral traits.

In either case, her case for global (as opposed to circumstantial) admiration would have nothing to do with natural duties. We have left the realm of the moral and entered the realm of the cool or trendy. We are of course entitled to admire and fawn over others for being cool, having sex appeal, or being rich as long as we do not claim that these are connected to duty. But this is only one type of differential respect. The other strategy indirectly ties respect to our status as beings with dignity. It is connected to our differing levels of performance of moral acts and will be the topic of the next subsection.

II.2.c Differential respect attached to dignity

An individual could argue that she has a moral claim to be shown a differential, non-circumstantial (that, universally applicable, at least within her broad community) level of respect in at least one of two ways. She could claim that she excels (or, if she is an honest scoundrel, is deficient) with respect to broad, morally relevant traits. The second, related reason is that she (1) has instantiated in good faith a set of moral rules and (2) is following those rules better than others are. She is actually fulfilling her duties as she understands them, including the more idiosyncratic duties that stem from more general ones.

It helps to think of this second reason in terms of Kant’s imperfect duties. Assuming I have an imperfect duty to provide charity to those who need it, I can discharge the duty in any number of ways. There are four considerations worth noting
here. First, as I think through my general duty of charity, I may decide that some ways of giving are better than others. For example, giving money is probably better than handing out free copies of this dissertation. Second, thinking through my general duty will cause me to generate subsidiary duties, i.e., “I ought to give away 10% of my gross income” or “Because I am good with people, I will personally give money and interact with those in need instead of donating it to foundations.” Third, the subsidiary duties would be unique to me. My particular circumstances would inform the shape and content of these duties. And finally, the subsidiary duties are not morally arbitrary; while recognizing that good-faith individuals will rightly generate different subsidiary duties, I will still judge my own moral performance as well as the moral performance of others. When judging others, I will be judging both the subsidiary duties they selected as well as their adherence to the duties that I think are correct (for them).

On this view, while everyone has dignity, those who discharge their moral duties deserve a type of respect that should be withheld from those who fail to uphold their duties. Dignity may provide a floor for how we may treat scoundrels, but it seems wrong to say that we should treat moral saints no differently than moral devils. Shouldn't going beyond one's duty of beneficence (assuming such a duty exists) and giving until it hurts merit some form of recognition and additional respect? And unlike the chess champion who demands to be admired for his chess ability or the non-moral traits that the chess ability implies (such as excellent memory or pattern recognition), those who are morally fastidious should be entitled to the recognition that they are not
mere people, but good people. \(^{298}\) In fact, it may be a moral requirement to show this recognition, however perfunctorily, when these good people are encountered.

What’s more, as I said earlier, people who take their duties seriously – and this includes philosophical anarchists – will, if they start from the premise that human beings have dignity, not only generate a set of rules that informs how they and others should be treated, but they will also generate a set of criteria to be used to evaluate and rank others in terms of moral stature. Naturally, these criteria will differ – I have already pointed out that there are differences between Kant, Wolff (a pseudo-Kantian), and Simmons (a pseudo-Lockean). As a result, people will differ when evaluating who counts as a good person. All may agree that people deserve differing levels of respect, but we will disagree as to who deserves more or less respect relative to their peers. \(^{299}\)

Here I want to anticipate and defuse a predictable objection. Some might say that respect should be given to people who sincerely follow their own doctrine, regardless of its content. If I were to adopt this position I would say that while I may not agree with Jane's expansive conception of right and wrong (even if we agree on the basic duties we owe others qua beings with dignity), as long as she sincerely follows her beliefs, she is necessarily entitled to significant differential respect – i.e., that I should treat her as morally superior to those who are laxer with respect to their

\(^{298}\) Igor Primoratz, drawing on Hegel, argues that even those who break the law for the sake of personal gain should be recognized – and punished – not just for objective reasons, but also out a recognition that the punishment would be subjectively acceptable to the perpetrator himself qua rational being. In other words: there is even a case for differentiation on behalf of the losers. See Primoratz, Justifying Legal Punishment, 77-78.

\(^{299}\) Again, I want to point out that this has no bearing on the baseline respect that we are all entitled to by virtue of our being creatures with dignity.
moral commitments. On this view, I may think how she fulfills a duty that we agree
she has, e.g., beneficence, is stupid or ineffective, but what counts is her willingness to
do the right thing as she understands it.

There are two arguments against this. The first objection is that there is no
prima facie reason to believe that sincerity in one’s beliefs is itself a reason for the
respect that I am currently discussing. This type of sincerity and steadfastness can be
seen as moral integrity, but it could also be seen as fanaticism. What’s more, even if
we do feel that such an attitude is worthy of respect, it is not clear (to me, in any case)
if it is admirable as a moral quality or simply a desirable non-moral trait. In other
words, I do not know if I would want that trait for myself because it would make me a
better person or because I enjoy the prospect of living with less self-doubt and think I
would be better off if I were more committed to my beliefs. If the former, it would
indeed be worthy of the type of moral respect I am discussing in this subsection. If
not, it may garner admiration, but not moral respect.

Second, and more importantly, the content of the set of sincerely held beliefs
can be repugnant. Many people believe that sincere terrorists are sincerely terrible.
Note that this can co-exist uncomfortably with admiration for the other traits
demonstrated by the people with the repugnant beliefs. Comedian and talk show host
Bill Maher lost his show, Politically Incorrect with Bill Maher, when he waded into
this ambiguous territory shortly after the 9/11 attacks. “We have been the cowards,
lobbing cruise missiles from 2000 miles away. That’s cowardly. Standing on the
airplane when it hits the building… say what you want about it; not cowardly.”300 The backlash to these remarks was predictably swift and livid, and the show was soon cancelled.

Maher’s problem was that he was unclear as to whether he was complimenting a trait that was not morally salient, the equivalent of saying “Terrorists sure are good at chess!” or complimenting a trait that is morally salient, thus implicitly bringing into question the moral condemnation of the terrorists. If the former, then his blunder consisted in suddenly discussing non-moral considerations in the middle of a clearly moralized discourse. If the latter, he needed to either explain how the terrorists could be just as morally condemnable while possessing morally desirable traits or to bite the bullet and explain why courageous terrorists are morally superior to cowardly ones. It seems far easier to say, alongside Kant, that traits such as courage are not morally salient.301 In any case, someone who says otherwise would need to have an answer to Bill Maher’s terrorist problem.

The difference between terrorists and those with whom we normally disagree on matters of moral importance is that terrorism is clearly so bad that we feel as though it should be vigorously resisted. Yet even those people whose imperfect beliefs do not need to be resisted per se can be ranked according to their moral behavior. We naturally feel that those whose actions coincide with our moral beliefs should be held in higher general esteem than people who mostly engage in what we consider to be frivolous or mildly immoral activities at the expense of discharging other moral duties.

300 Sims, “Tragedy + Comedy.”
301 Kant, Metaphysics of Morals, 7-8.
The upshot is that we will (morally) judge others in part based on their adherence to rules that emerge from our sincerely held beliefs, even if we do so silently. We may treat everyone respectfully in some sense because we have equal dignity, but we will demonstrate more respect (in some sense) to others, and we will likely feel that demonstrating this additional respect to some – and relative disrespect to others – is appropriate. What’s more, each of us, being a person with dignity, will come to expect that others should treat us in a certain way. If I feel that I should treat someone who adheres to my sincere moral beliefs (which, to repeat, I must generate as per my status as a being with dignity) in a respectful way, then I must demand that others treat me that way.

III. Natural duties and deliberation

It is time to pull together the threads of this chapter into a position that responds to the philosophical anarchist challenge. So far, I proposed a model that differentiates dignity and respect. Beings that possess dignity are entitled to a certain type of respect. This respect is universal in that it applies unconditionally to all beings with dignity. It is also egalitarian since it prescribes the same duties to all possessors of dignity. Call this type of respect "universal respect."

I have also suggested that some forms of respect are not directly tied to dignity, at least not in the same way as universal respect. Some forms of respect are tied to considerations that we do not consider morally salient, such as skill at chess or football. These are irrelevant to the question of political obligation as far as
philosophical anarchists are concerned. But another type of respect is relevant. It's the kind of respect that we feel we owe those who consistently perform in a morally upright manner, and the kind of respect we withhold from those who consistently shirk their moral duties or who adopt and live by what we consider morally defective frameworks. Our respect differs from person to person based on our evaluative criteria, hence why I call it "differential respect."

In other words, universal respect brings us together, and differential respect pushes us apart. Charles Larmore has argued that "Others are due equal respect by virtue of their capacity for working out a coherent view of the world and indeed the good life," but even coherent views that are not obviously wrong are not necessarily worthy of respect. We may withhold a certain kind of respect – and ignore the duties that that type of respect brings about – when faced with such people. I believe that this is the central tension that earnest philosophical anarchists must face, and this especially with respect to a duty to deliberate. I will discuss deliberation more fully in the next chapter, but here I will sketch out deliberation’s connection with natural duty, and why this matters for philosophical anarchists.

Our recognition of others as beings with dignity gives rise to a duty to attempt to deliberate with them. Someone like Darwall might say that deliberation is part and parcel of demonstrating the respect one has for a being with dignity. If Darwall is right, deliberation is clearly tied to the essence of dignity and the respect we show others that stem from this dignity. It would mean that treating people as people

requires addressing them while recognizing and respecting their capacity their ability to make autonomous judgments. Not deliberating or deliberately deliberating disrespectfully entails an implicit contradiction; it would mean treating people contrary to the status that we know they have.

But even without recourse to Darwall, we can still say that we owe a duty of deliberation to those with which may come into contact simply by virtue of living in a world of moderate scarcity. Proximity counts, and we owe it to others to try to agree on common rules of engagement. Wellman makes something close to this point when he says that "harmful chaos would prevail unless virtually all of those in spatial proximity defer to the same set of clearly defined rules." While I disagree that we necessarily need the same set of rules, we are faced with a stark choice: try to come to terms with others regarding “political” issues or predictably violate our natural duties by inadvertently running roughshod over the prepolitical rights of others. As a result, we seem to have an actual duty to deliberate.

Happily for archists, this could go some way to answering the philosophical anarchist objection to natural duty. We could have a special bond duty to “our” government if it is the institution that is needed to help us discharge our duties to those who are close to us. The difference between my support for the US government and the Swedish one is that without the former, I’m very likely to inadvertently violate the natural duties of others as I go about my daily life. If this model holds, then Simmons’ objection to natural duty theories is overcome: we have a natural duty to support just

---

303 Wellman and Simmons, *Duty to Obey*, 38.
institutions (which Simmons concedes), and here we have an argument explaining why we have a duty to talk with others in order to discuss and maintain our institutions. To fail to do so would mean to fail to respect those around us.

But this duty to deliberate is not absolute. Part of our duty to articulate fundamental duties (e.g., the duty to beneficence) into particular duties (e.g., the duty that I believe I have to donate a percentage of my gross income to certain types of charities) includes the duty to determine what kinds of deliberation count as respectful and which count as disrespectful. Some duties will (apparently) be universally applicable. Others will be circumstantial. Crucially, the formulation of these particular duties includes a duty to decide what rules to apply to the utterances of our interlocutors in order to determine if they are being respectful to us. To the extent they are not being respectful, our duty to engage in deliberation has been discharged. To put it simply: we do not have a duty to engage in deliberation with those who use deliberation to disrespect us.

This chapter has not provided a direct archist defense against the philosophical anarchist, but it has provided a series of distinctions which will ground my investigation of deliberation as a possible response to the philosophical anarchist challenge. To jump ahead in my argument briefly: respect must be demonstrated via how we engage in discourse, and that the combination of our different moral perspectives and the many ways that we can fail to demonstrate respect by way of

---

304 In fact, I will cast doubt on this in chapter 7. To foreshadow the point: even apparently obvious rules like “no ad hominem attacks” can fail to apply in some cases, leaving us with overly vague “rules” that require further specification (which we will only be able to do on a case-by-case basis).
differing communicative styles will cause delegitimizing communicative breakdowns.

To sum up, existing natural duty theories fail to satisfy philosophical anarchists because they have difficulty tying us to “our” state. Yet in spite of this, I believe that natural duty is the most promising ground of political obligation. If we accept that we have equal dignity, then that dignity provides a floor for how poorly we can treat each other. Our status as beings with dignity entitles us to a certain type of equal respect. We can also claim to be entitled to more or less respect, however, based on our possessing desirable traits (which should not entitled us to a generalizable respect) or on our fulfilling moral obligations that stem from our understanding of our status as beings with dignity (which should entitle us to a generalizable respect). This type of respect cannot be communicated in the abstract; it must be demonstrated in discourse itself, by way of communicative structures such as politeness, etiquette, and apology. I will turn to those in a later chapter, but first I will discuss deliberative democracy more fully in the next chapter.
Chapter 6: Deliberative Democracy

I. Why Deliberative Democracy?

Although the literature on deliberative democracy has mushroomed since John Bessette coined the term in 1980, it has not been suggested as a serious answer to philosophical anarchists.\(^{305}\) This seems odd. After all, deliberation is undoubtedly somehow connected to legitimacy.

Deliberation's absence in the philosophical anarchism debate makes sense when we consider that unlike happiness or justice, virtually no one prizes deliberation for its own sake. When it is discussed in the context of legitimation it usually ends up being a second-order theory, one that gives heft to one of the first-order legitimation theories described above. For instance, some say that deliberation yields outcomes that have greater utility.\(^{306}\) Still others – and I place myself in this camp – may say that we cannot discharge certain natural duties absent deliberation. Deliberation itself, however, is usually not used to argue for or against state legitimacy.

Explicitly connecting deliberation to debates on political legitimation is further complicated by the fact that deliberation is frequently discussed in a context which presupposes the existence of a more or less legitimate state. No one says that the real problem with a plainly illegitimate modern state is a lack of deliberation. To return to the three questions I discussed in Chapter 1, theories of deliberation have assumed that

\(^{305}\) Bessette, “Deliberative Democracy.”
\(^{306}\) Mill, *Representative Government.*
a legitimate state is possible and that some states are at least minimally legitimate.\textsuperscript{307}

In other words, they address my third question, presenting ways that legitimate states can improve, for instance, by making government more responsive to the will of citizens, reducing levels of unfairness (even if the existing level is sufficient for the state to be legitimate), or by increasing feelings of solidarity. I will address these in more detail later in the chapter; for now I want to emphasize the assumption of legitimacy that these theories tend to make. For instance, Rawls' recent theory relies extensively on deliberation, yet does so after taking for granted the existence of a more or less well defined polity, one where the question of the existence of a meaningful political community has been settled.\textsuperscript{308} The question that I am looking to answer is whether or not deliberation can help us answer philosophical anarchists without presupposing an answer to one of their most fundamental questions, i.e., whether or not people belong to a polity in a morally meaningful sense.

My aim in this chapter is thus quite limited. The idea is to provide a history of and commentary on important strands of the deliberative democracy literature, not to add to it on its own terms nor to provide an exegesis for its own sake. I am discussing themes, not comprehensive accounts. My aim is to map a framework that will allow me to argue for the importance of deliberation in a story that purports to answer the concerns of philosophical anarchists. Ultimately, in this chapter I am describing deliberative processes, and challenges to these processes, at a high level. The next two

\begin{flushright}
\textsuperscript{307} As a reminder, the questions are: (1) could the state be legitimate? (2) if a state could be legitimate, what other criteria must be met for a particular state to be legitimate? and (3) if a state were legitimate, what kinds of institutions, policies, or other considerations should we endorse? \\
\textsuperscript{308} For this critique of Rawls, see for instance Carter, “Political Liberalism,” 141-145.
\end{flushright}
chapters will discuss micro-level considerations which stem from the account of deliberation that I will present here.

I have organized the chapter as follows. In the first section, I present some of the objections to non-deliberative liberal democracies that motivate the deliberative democracy project. I then describe standard arguments for deliberative democracy. Next, I describe the deliberative democratic project, broadly construed, along with ostensible benefits of deliberation. In the second section, I describe "orthodox" deliberative theories as well as the "subaltern" theories that aim to improve on them.

In the third section, I critique the subaltern theories. Finally, in the fourth section I describe, in broad strokes, a deliberative model that I believe has the best chance of adequately answering the philosophical anarchist challenge. The thrust of the model is that, as I described in the previous chapter, we may have a duty to deliberate, but each (good-faith) person must determine for herself whether or not participation in a particular deliberative forum or practice is consistent with her self-respect.

I.1 Problems with democracy

Deliberative democracy is a family of theories that seek to address some troubling aspects of modern liberal democracies. If the deliberative democracy literature is sprawling, it is because the range of problems it identifies and tries to remedy is far-reaching. The following are some of the main concerns that deliberative democrats have tried to address.
I.1.a Coping with pluralism

First, deliberative democrats take seriously what Rawls has termed the fact of reasonable pluralism. How are we to govern ourselves according to democratic ideals in contexts where there are deep social divisions based on reasonable disagreements on fundamental political, ethical, and metaphysical matters? Democratic theories that try to skate around this problem by ignoring it or suggesting a set of common, stated values that can be used as a touchstone to adjudicate between differences are, deliberative democrats argue, understating just how fundamental the disagreements can be.

Indeed, it is this pluralism that caused Rawls to abandon his attempt to ground a far-ranging political theory on an interpretation of Kantianism. The problem is not so much that his Kant-flavored Theory of Justice is untenable, but rather, as Rawls recognized, that others could reasonably disagree with its Kantian foundations. A Theory of Justice – or any other theory based on a reasonable but controversial comprehensive doctrine – could only hold if one of three conditions obtained: (1) that the polity were sufficiently homogeneous for the citizens to agree on fundamental philosophical questions; (2) that there were no reasonable alternative answers to basic philosophical questions or, to put it another way, that no reasonable person could adopt an alternative moral framework; or (3) that there were some kind of meta-justificatory method that would allow a theory such as that found in A Theory of

309 Rawls, Political Liberalism, xvii.
310 Ibid., xv-xvii.
Justice to be one of many plausible justificatory theories.

Recognizing the fact of pluralism involves, at its most basic, the rejection of (1). There is an insufficient level of agreement on fundamental moral frameworks in modern societies and it is unlikely that this situation will be reversed any time soon. Republicans might disagree here, either by downplaying the level of pluralism or by proposing to redouble efforts to bind citizens by encouraging them to participate in shared projects or by socializing them into strongly identifying with a national symbol or history.

Both of these responses are problematic. Anyone who has grown up in a multilingual, multiethnic city can immediately attest to the fact of pluralism. What’s more, trying to significantly reduce the existing level of pluralism seems quixotic in the face of what appears to be a trend towards increasing amount of social fragmentation and polarization. Finally, creating solidarity by an appeal to a feature such as a shared narrative or nationalistic identity seems like a partial solution at best. Even if this project were somewhat successful, it is unlikely that this nascent common identity would trump longstanding viewpoints during times of moral crisis. Such a solution barely seems like an improvement over what Rawls calls a modus vivendi - a fragile truce between players with different aims and interests, one that either player would violate if it thought that it was in its interest to do so.\footnote{Ibid., 147.}

This concern over the creation of a mere modus vivendi might lead us to endorse (2), that the fact of pluralism forces us to abandon the hope of constructing
and maintaining a meaningful political community. In this conception, we abandon the possibility of overcoming our alienation towards those whose worldviews differ significantly from ours. Politics here amounts to bargaining, a notion that proves depressing to most democratic theorists. Not only does it fly in the face of ineffable democratic ideals, but is it also an unstable basis for social order and therefore problematic on practical grounds. I should add that it would also force archists to abandon the hope that deliberation could provide the foundation of an answer to philosophical anarchists.

Given these worries, deliberative democrats propose (3), a meta-theory that would provide sufficient political glue for a society that is stable for the right reasons while allowing participants to keep their worldview. One idea is to use deliberation to allow people with different fundamental perspectives to come together and make agreements that are acceptable to each from his point of view. Another idea – one which will play a central role in the modified theory that I will endorse – is that deliberation allows participants to demonstrate respect to each other.

I.1.b Other problems with democracy

312 To be fair, this "bargaining" that takes place isn't necessarily the kind that occurs between two basically materialist people trying to get as much stuff as they can. In fact, that would be unlikely on this reading of the fact of pluralism, since it is the lack of homogeneity of outlooks that spurs the modus vivendi. When we think "bargaining" here, we should think not of car dealers and buyers trying to squeeze every dollar out of the lemon, but of a monk, a goth punk, and an inveterate fashion hound trying to get along.

313 Rawls, Political Liberalism, 147.

314 This innocuous way of putting it obscures a fundamental question that deliberative democrats must face: does there need to be a shared perspective or can participants agree for reasons that pertain to their own, idiosyncratic beliefs? I will address this question below, but the long and short of it is that whereas most deliberative democrats endorse the former, I prefer the latter.
Besides having a healthy respect for the fact of pluralism, deliberative democrats are also keen to reject theories of democracy that embrace *elitism*. Elitism has been justified on plausible grounds, including (1) anxieties surrounding the emergence of mass involvement in electoral politics\(^{315}\), (2) institutional arguments against electoral procedures, i.e., that they cannot locate the existence of a general will and/or cannot translate such a(n unidentifiable) will into policy, and that this leads to a “gaming” of the system by those who can engage in institutional manipulations\(^{316}\), and (3) epistemic grounds, stretching all the way back to Plato's arguments for philosopher-kings in *The Republic*. These objections, while at least superficially plausible, are nonetheless troubling for people committed to democratic ideals. Deliberative democrats provide one line of argument against elitism, arguing proper deliberation allows citizens to participate in politics in a productive way, to form (rather than discover) a common will, and generate policies that tend to be of relatively high quality.

Third, some deliberative democrats argue that aggregative conceptions of democracy – those that focus on atomized voters and which form the basis of much modern political science – are flawed, mainly because they inappropriately normalize and stabilize both individuals' preferences and the institutional setting.\(^{317}\) In other words, aggregative theories of democracy paint a portrait of agency that may be too


\(^{316}\) See Riker, *Liberalism Against Populism*.

inflexible. These models take preferences not only as given but also as the conceptual starting point of analysis. This means that they tend to ignore the distributions of power that lead to these preferences. Let's accept that an aggregative model may be equipped to translate actual preferences into commensurate outcomes (William Riker notwithstanding).\textsuperscript{318} Such a model can account for Joe's desire for higher welfare payments but would be ill-equipped to decide between antecedent institutional patterns that inform Joe's preferences in the first place. As Amy Gutmann & Dennis Thompson point out, not only are some methods of aggregation are more appropriate in some cases than in others, but aggregative models suffer the flaw of not allowing citizens to challenge the aggregative measures themselves.\textsuperscript{319}

Fourth, deliberative democrats frequently argue that the status quo – say, in 2017 America – does not allow for the expression of a clear public voice. Public opinion polls, talk radio, meeting with representatives, and lobbying, while in some sense allowing for the expression of public sentiments, are ultimately insufficient to allow for meaningful participation and expression of preferences. As John Gastil says,

\begin{quote}
In practice, policymakers often act without knowing how the general public might view an issue, and those citizens who find fault with such actions routinely state their preferences after the fact by criticizing officials for the actions they have taken. On a given issue, an elected official might hear the public's voice as a warning, a request, praise, criticism, or merely a notice. Under these circumstances, it is not surprising that officials respond to lobbyists, who underscore the real consequences of following their advice and offer a measure of genuine policy expertise. The situation might be different if the public could express itself in a way that was representative, deliberative, and articulate.\textsuperscript{320}
\end{quote}

\begin{flushleft}
\textsuperscript{318} For a thorough critique of Riker, see Mackie, Democracy Defended. \\
\textsuperscript{319} Gutmann and Thompson, Why Deliberative Democracy? 13-16. \\
\textsuperscript{320} Gastil, Popular Demand, 111.
\end{flushleft}
Whether or not anything approaching a public voice or general will exists, those who translate preferences into policy currently have a tough time identifying and knowing what to do with signals in a sea of noise. Deliberation, the argument goes, would help policymakers identify, clarify, and better interpret their constituents' preferences.

1.2 Benefits of Deliberation

Deliberative democrats have collectively provided a number of positive justifications for deliberative democracy. Some of these directly try to remedy the ostensible problems of liberal democracy, while others cite additional, positive benefits of deliberation.

1.3.a Better outcomes

First, deliberation may provide direct epistemic benefits to the political process. The idea is that legitimacy is at least tangentially enhanced when the political system comes close – or close to as close as possible – to discerning and acting on the truth.321 These types of arguments for deliberation are not new; John Stuart Mill argued that a properly organized parliament should spend the vast majority of its time deliberating rather than designing or implementing laws.322 Nor must deliberation be restricted to parliamentarians to be effective; modern deliberative democrats mainly focus on citizen-to-elite or citizen-to-citizen deliberation. Assuming that there is some

321 Pincione and Tesón, Rational Choice, 3.
322 Mill, Representative Government, 111-114.
substantively correct answer to a political question – that is, one in which the truth value of the answer does not wholly depend on its emerging from a particular decision procedure – deliberative democracy may get us closer to this truth by mixing democracy with expertise.

Mark Warren lists a number of reasons for this. First, deliberative procedures may reduce the amount of information that actors strategically withhold. In a hierarchical structure, those who have access to expert opinions are likely to "hoard" them, a prospect that is mitigated in a system that promotes a free flow of information and opinion. Second, deliberation allows input from people who have practical, on-the-ground knowledge or expertise – techné – but who normally would not have much input in decision formation. Finally, experts become more trustworthy precisely because they (and their opinions) become more accessible to the public. According to

324 The value of this kind of bottom-up knowledge can be seen in the business world, where there are relatively clear ways (i.e.: profit over time) of gauging the quality of ideas. Charles Duhigg provides an example of this in The Power of Habit, 117-118:

“"I [O’Neill, the new CEO of Alcoa] said to the hourly workers, ‘If your management doesn’t follow up on safety issues, then call me at home, here’s my number,’” O’Neill told me. “Workers started calling, but they didn’t want to talk about accidents. They wanted to talk about all these other great ideas.”

“"The Alcoa plant that manufactured aluminum siding for houses, for instance, had been struggling for years because executives would try to anticipate popular colors and inevitably guess wrong. They would pay consultants millions of dollars to choose shades of paint and six months later, the warehouse would be overflowing with “sunburst yellow” and out of suddenly in-demand “hunter green.” One day, a low-level employee made a suggestion that quickly worked its way to the general manager: If they grouped all the painting machines together, they could switch out the pigments faster and become more nimble in responding to shifts in customer demand. Within a year, profits on aluminum siding doubled.”

“It turns out this guy had been suggesting this painting idea for a decade but hadn't told anyone in management,” an Alcoa executive told me. "Then he figures, since we keep on asking for safety recommendations, why not tell them about this other idea? It was like he gave us the winning lottery numbers.”"
Warren, experts cannot hide behind their position or status; citizens can, over time, begin to detect which experts to trust and which are frauds or incompetents.

What's more, there is some evidence that even in the absence of experts, small group communication can yield better decisions, including technical decisions, compared to decisions made by individuals.\textsuperscript{325} The most plausible explanation for this is that by sharing information, those with erroneous beliefs will come to recognize the veracity of the opposing argument and will modify their opinions. This usually amounts to a modification of beliefs concerning \textit{means} – i.e., “I suppose disarmament is more conducive to peace than my initial idea of torturing enemy captives” – but it can also cause individuals to re-assess some goals.

What's more, as Diego Gambetta points out, deliberation not allows the sharing of information, but also helps the participants to find novel solutions to problems: “Deliberation, in addition, spurs the imagination indirectly if it reveals that, on all known options, no compromise is possible, for this provides an incentive to think of new ones. By the same process, it can instill the courage necessary to embrace solutions that were thought to be too daring before it became clear that no compromise was otherwise possible.”\textsuperscript{326} Deliberation can clarify when the obvious solutions won't work, and as a result, the participants realize, along with Sherlock Holmes, that "we must fall back upon the old axiom that when all other contingencies fail, whatever remains, however improbable, must be the truth."\textsuperscript{327} Deliberation can thus give a

\textsuperscript{325} Gastil, \textit{Popular Demand}, 23.
\textsuperscript{326} Gambetta, "‘Claro!’" 22-23.
\textsuperscript{327} Doyle, "Bruce-Partington Plans."
relatively sober group the push to adopt uncharacteristically bold policies.

The change in beliefs may potentially go deeper than this. One of the major appeals of deliberation is its potential *transformative* effects on the actual preferences of the deliberators. Rather than change my opinion on the basis of means-ends rationality or as part of a strategic compromise, I may come to revise my underlying preferences. A hawk may, after sufficient discussion, undergo a shift in values and become a dove. In other words, deliberative democrats believe that we should assume that there are no permanently fixed preferences and that all beliefs are subject to revision. Deliberative democrats thus have a way of dealing with the "fact of pluralism" that is not available to proponents of strict aggregative democracy. To the extent that preferences can change and converge, deliberation reduces the number (and hopefully the intensity) of disagreements.

Of course, change can be hard. Some preferences may only change in clusters as an individual adopts a whole new intellectual paradigm. Gerry Mackie has argued that "a given desire or belief is not isolated, but, rather, is located in a network structure of attitudes, such that that persuasion sufficient to change an attitude in isolation is not sufficient to change the attitude as supported by its network." In other words: deliberation may change minds, but probably not as much or as often as deliberative democrats would hope. Still, the importance of this type of shift should not be understated – to the extent that it is possible to align our underlying preferences, we need not worry about deep social divisions or institutional problems of

---

329 Although one consequence of this view is that many changes may occur simultaneously.
aggregating votes. The transformative effects of deliberation seem to give us a better chance of addressing deep moral conflict.\textsuperscript{330}

What’s more, the disappointment deliberative democrats feel in the face of an unimpressive transformational occurrence may be mitigated depending on why exactly they think the transformation matters. In \textit{The Market and the Forum}, Jon Elster advances seven objections that cast a shadow over the contention that transformation occurs as a result of participants’ being forced, by the nature of deliberation, to advance their interests in terms that are not nakedly self-serving.\textsuperscript{331} After casting doubt on whether or not deliberation can achieve its goals of unanimity and whether or not the results of the transformational experience would be desirable, Elster addresses the proposition that couching arguments in terms of the common good will “purge the desires of all selfish arguments.”\textsuperscript{332} He argues that any attempt to realize the common good – understood not necessarily as the instantiation of an ideal but as an arrangement that is preferably to one where individuals nakedly try to enact personal preferences – will benefit one or more specific groups unduly, who will then have good reason to insist on one arrangement over another. These preferences can always be justified in terms of common good, even if the group members are clearly holding fast to private incentives. In sum, here Elster provides an additional argument against the contention that deliberation is transformative, if the transformation is meant to refer to some kind of a deep feeling of identification. Deliberative democrats who are

\textsuperscript{332} Ibid., 17.
after this “deep” transformation may in fact be out of luck.

Yet Elster casts a different kind of doubt in, among other places, his later
*Deliberation and Constitution Making*. Here he gives reasons why someone may want
to misrepresent his preferences in a deliberative setting, and in doing so provides a
reason to believe that deliberation can lead to a transformation by providing
constraints that consistently modify the expression of personal preferences. He is
worth quoting at length:

> The main argument I have tried to make is that a *deliberative setting*
can shape outcomes independently of the motives of the participants.
Because there are powerful norms against naked appeals to interest or
prejudice, speakers have to justify their proposals by the public interest.
Because there are powerful norms against the use of threats, they have
to disguise them as warnings. Moreover – and this is the key point from
the behavior point of view – the proposals will be *modified as well as disguised*. The disguise of private interest or prejudice is subject to two
constraints. First, as I mentioned, there is the *imperfection constraint*.
Because a perfect coincidence between private interest or prejudice and
impartial argument is suspicious, self-interested or prejudiced speakers
have an incentive to argue for a position that differs somewhat from
their ideal point. Second, public speaking is subject to a *consistency constraint*. Once a speaker has adopted an impartial argument because
it corresponds to his interest or prejudice, he will be seen as
opportunistic if he deviates from it when it ceases to serve his needs.333

Here transformation really does occur if it is understood to refer to the transformation
of either (1) the policies that are actually being put forth and acted on, and/or (2) the
*persona* of the speaker. Whether or not this transformation is necessarily good is up
for debate334, but deliberative democrats can, if worried when faced with arguments
casting doubt on arguments that suggest that deliberation is deeply transformative,

---

334 Elster provides reasons to worry about the transformational effects of deliberation in “The Market
and the Forum.”
switch to a definition of transformation which may be satisfactory and leave them on surer ground. Yes, groups may still seek their private interest, but to the extent that the consistency constraint applies, the justifications they offer may bind them to unfavorable policies when the circumstances change. For instance, if proponents of the coal industry have historically justified requests for a favorable regulatory environment partially on the grounds that coal mining provided national energy independence, the same logic may leave them speechless when circumstances show that investment in renewable energy would better achieve those aims.

I.3.b Respect

Deliberation's transformative effects bridge the gap between those who believe that deliberation is instrumentally valuable and those who believe that it either has intrinsic worth. I am interested in these non-instrumental versions of deliberation. It's not that we should ignore the epistemic virtues and resultant good outcomes of deliberation; most people, including philosophical anarchists, would prefer a procedure that produces better outcomes over one that yields worse outcomes. The problem is one of relevance. To use the terms I laid out in Chapter 1: epistemic considerations tend to be tied to the third question about legitimacy – Given a

335 I recognize that this is not as self-evident as it first seems. At least two exceptions come to mind. First, some (usually leftist) activists complain that incrementalistic politics pacify would-be revolutionaries and forestall radical change. In other words, we are collectively giving up a jackpot in order to win small pots. Second, positive outcomes that come about the wrong way can lead to negative externalities. William Godwin and John Stuart Mill both worry that policies that formulated and implemented without citizen engagement dull the intellect and initiative of the people, resulting in an apathetic and intellectually impotent citizenry. Clark, *Philosophical Anarchism*, Chapter 8; Mill, *Representative Government*, 48-53.
politically legitimate system, how should we govern ourselves? – instead of the second
– What does it take for a system to be legitimate? The transformative effects of
deliberation go beyond yielding better outcomes; they also may legitimize in at least
one of two possible ways.

The first and most ambitious way that deliberation could legitimize is by
transforming a motley group of deliberators into citizens that hold deeply shared
values. This type of solution would please republicans. While this transformation is
not impossible, I will summarily dismiss it for two reasons. First, this view seems to
simply reject the idea of "the fact of pluralism" that motivated the recent interest in
deliberation in the first place. If a deeply diverse society can come together and create
a meaningful shared identity through deliberation, it is hard to imagine that the
factions did not share deeply similar values in the first place. When pluralism is
attached to deep questions of identity, it may be possible to harmonize interests, but it
is difficult to supersede or replace them.336 Second, the worries that I express in the
subsequent chapters of this dissertation – that it is difficult for diversity-embracing
liberals (who are sympathetic to deliberative democracy) to deliberate in a legitimizing
way – apply even more to this republican view. Mass support of any particular
worldview will make it difficult for citizens to understand those oddballs whose
perspective differs from the norm.

The second way that deliberation could legitimize is by convincing participants
that the deliberative procedure takes them seriously and shows them respect. For

instance, Robert Goodin argues that “[…] [deliberative] democracy […] is essentially a matter of treating […] autonomous individuals with equal consideration and respect.” It is this, and not necessarily the outcomes themselves (although the outcomes might have to satisfy some minimal threshold), that legitimize. What this respect consists of is, of course, up for debate.

I believe that this is this most promising legitimizing strategy. If philosophical anarchists could be convinced that (1) they had a duty to enter deliberations, and (2) their interlocutors treated them with the respect befitting an autonomous individual, then they may have a duty to obey the law, since the law would be the ultimate outcome of a collective decision making procedure in which they took part. My main worry - and this concern will inform the remainder of this dissertation - lies in how this respect is specified. It is not enough to have respect; one must show it and show it well. This will oftentimes translate to "show it to the satisfaction of the person who is earnestly interacting with you."

I.3 Deliberation Defined

I.3.a Attempts at a Definition

Virtually all democracies are deliberative to some degree. And although nearly all democratic theorists agree that deliberation can bring about side-benefits – more on that below – for my purposes I want to focus on those theories which, as Chantal Mouffe says, “... claim that it is possible, thanks to adequate procedures of

---

338 In fact, it is at times arguably impossible to have it without showing it.
deliberation, to reach forms of agreement that would satisfy both rationality (understood as defence of liberal rights) and democratic legitimacy (as represented by popular sovereignty).” These types of deliberative democracy thus say that deliberation is central to democratic legitimacy. Deliberation is not merely a helpful feature, like air conditioning in a car. Without deliberation, the proverbial car falls apart.

Gutmann and Thompson have suggested defining deliberative democracy according to four criteria. First, a proper theory asserts that those who make policy decisions must support those decisions with reasons. Second, it must do so by appealing to principles “... that individuals who are trying to find fair terms of cooperation cannot reasonably reject.” Third, the deliberation must produce binding decisions. Fourth, the process, while producing periodic binding results, must be ongoing. Questions may be adjudicated provisionally, but nothing need be settled once and for all, as citizens are permitted to re-open previously settled issues in future deliberative proceedings. Put together, we have this definition of deliberative democracy:

A form of government in which free and equal citizens (and their representatives) justify decisions in a process in which they give one or another reasons that are mutually acceptable and accessible, with the aim of reaching conclusions that are binding in the present on all citizens but open to challenge in the future.

I should note the inclusion of “all citizens” in this definition. Deliberative democracy

339  Mouffe, The Democratic Paradox, 83.
341  Ibid., 7.
is not to be binding merely on deliberators. Given universal suffrage, it must also be binding on those who do not deliberate as well, much the same way as the outcome of a vote is usually taken to be binding on those who abstain.

Gutmann and Thompson believe deliberation has both instrumental and expressive value and that these values complement each other. A procedure's not producing instrumentally good outcomes would "hardly express mutual respect," and discounting the expressive function of deliberation makes it less likely that policymakers truly understand the interests of the citizenry. Gutmann and Thompson further argue that deliberative democracy, properly construed, satisfies both proceduralist and substantive conceptions of democracy. That deliberation has something to do with procedural questions is obvious. And while there are those who may say that imposing any substantive principles on deliberation constrains democratic decision-making, Gutmann and Thompson say that procedures must at some level be justified by an appeal to substance. In other words, no decision procedure justifies itself. In this case, the procedures legitimize because they can be defended in terms of the antecedent principle of, among others, fairness. I, on the other hand, am ultimately trying to justify it in terms of natural duty.

While Gutmann and Thompson present an attractive and plausible account of

342 Ibid., 23.
343 Ibid., 25. Of course, they are not the only ones to do so. For more general statements on the importance of generating substantively good policy to bring about legitimate authority, see Elster, “Market and the Forum.” David Estlund proposes and defends a middle ground between "impartialist proceduralism" and epistemic theories that endorse the procedures most likely to get the "correct" substantive outcome, which he calls "epistemic proceduralism." See Estlund, “Beyond Fairness.”
deliberation, their definition also highlights the unavoidable difficulty in pinning down what deliberation actually looks like. Many frameworks could satisfy their four criteria. As Richard Posner says, “Any effort, my own included, to define or describe 'deliberative democracy' risks lumpiness because conceptions of deliberative democracy differ among a number of dimensions.”\(^{344}\) These dimensions include the "abstractness" of deliberation, the stringency of necessary conditions for deliberation to be legitimizing, the part of the political structure where deliberation is to take place, and the degree to which deliberation is seen as instrumentally vs. intrinsically valuable.\(^{345}\)

There are thus many possible theories of deliberation, which Philip Pettit tries to consolidate by arguing that they share three constraints. First is an inclusiveness constraint, whereby all members are "equally entitled to vote on how to resolve relevant collective issues […] with something less than a unanimous vote being sufficient to determine the outcome." Second is a judgmental constraint, which states that members should deliberate with reference to common concerns before voting. The third constraint is "dialogical"; the debate should take place "in open and unforced dialogue with one another whether in a centralized forum or various decentralized contexts."\(^{346}\)

Pettit’s attempt at consolidation, plausible as it is, once again shows the conceptual vagueness of deliberative democracy. Must deliberators *vote* for the

\(^{344}\) Posner, *Law, Pragmatism, and Democracy*, 133.  
\(^{345}\) Ibid., 133-135.  
\(^{346}\) Pettit, “Deliberative Democracy,” 139.
deliberation to count? Must this vote be direct, or can participants vote for representatives who are well situated to hear (and be electorally responsible to) a robust, inclusive citizen debate on important issues? Is it enough for members to deliberate on what they intuitively believe to be mutually shared grounds, or must they explicitly come to an agreement as to what types of arguments and issues are permissible?

I do not want to caricaturize these theories of deliberation and, to be fair, they attempt to provide valuable distinctions that give the reader answers to these types of questions. Pettit, for instance, in articulating his ideal of non-domination, develops a notion of “discursive control” which could in principle help us cast judgment on deliberative practices. Pettit argues that freedom consists of rational control, that our actions are directed by our desires, volitional control, that we be moved to identify with those desires, and discursive control, which situates rational and volitional control within the context of social relationships where one is recognized by others as a person capable of intersubjective influence that renders her unfree.\(^{347}\) As Patchen Markell explains, Pettit wants this latter to be understood as “refusing to subject someone to forms of influence that bypass her ability to engage in the activity of reasoning in common.”\(^{348}\) Fair enough; this can help guide us when describing the details of appropriate deliberative settings but Pettit’s idea of freedom can be reasonably contested. Markell, for instance, argues that Pettit’s own theory commits him to a view of freedom where the value of discursive control should be understood

\(^{347}\) Markell, “The Insufficiency of Non-Domination,” 16-17.
\(^{348}\) Ibid., 17.
not only as a mechanism to prevent abridgments to freedom as he understands it, but also to help individuals construct an identity which would give them the capacity to be free. If Markell’s argument is right, then deliberative fora should be judged in part according to how well they accomplish that goal as well.

There are two points I take away from this. The first, minor point is that existing theories of deliberation are routinely scrutinized and challenged at the level of abstraction at which they are presented. There is nothing wrong with this – a better tuned higher-order theory is better than a cruder one. But the second, more important point is that there remains a gap between the higher-order theories and the lower-order ones. Even if Markell is right, what does the resultant deliberative model actually look like? We may have more specificity with respect to the desired outcomes, but, as anyone who has worked with students closely knows, merely specifying outcomes leaves a gap between theory and practice. “Show respect and encourage others to speak” is analogous to the trite admonition instructors give their students to “Be more clear.” What appears to be concrete instruction is not only surprisingly unhelpful to those who are not already in the know as to how to perform respect, encouragement, or clarity, but it positively impedes these goals at times.

First, not only does it invite paternalism when we are confronted with someone who doesn’t “get it” – Ward seems well-meaning but doesn’t understand how to be respectful, so let’s show him – but also when idiosyncratic ways of demonstrating these values backfires. There is clearly a need to clarify this nitty-gritty level of

349 Ibid., 19-20.
communication; at the time of this writing, Gary Chapman’s *The Five Love Languages*, which has as its thesis the idea that failing to communicate affection in the way our partner “hears” love will produce discontent, despite our best intentions, has been on the New York Times best seller list 297 weeks.\textsuperscript{350} Incompetently trying to explain how to be respectful can lead to more hostility, a point I discuss at length in chapters 7 and 8. Second, we are still left with the awkward problem of knowing what counts as respectful behavior. There will be cases where communication is clear but still insulting. Theories of deliberation that operate at a high level tend to gloss over these issues.

In sum, most deliberative schemes are conceptually vague, both at the level of abstraction in which they operate (since they can be challenged by other deliberative theorists), as well as in terms of how they connect their schemes to actual deliberative practices. To foreshadow: in the next two chapters, I am especially interested in problematizing this latter concern. I am not saying that higher-order theories of deliberation are wrong, only that they are the first step in establishing a suitable model. There is value in providing criteria that can help us accept or reject on-the-ground institutions and practices. For deliberative democracy to be useful, however, it can't remain at such a high level of abstraction any more than I can resolve long-standing ethical debates by saying that "we should treat people with respect," full stop. It is this lack of specificity that has motivated me to write the next two chapters, where

\textsuperscript{350} Chapman, *The Five Love Languages*. This footnote was written April 1, 2017. For the record, Chapman’s five love languages are: words of affirmation, quality time, receiving gifts, acts of service, and physical touch.
I raise challenges to nebulous theories of deliberation.

I.3.b Contrived vs. diffuse deliberation

To be fair, some have specified what on-the-ground deliberation should look like (even if they often do not take into account questions of etiquette, which I address explicitly in the next chapter), and, unsurprisingly, the proposals vary tremendously in their particulars. As Posner says, the deliberative schema can differ on any number of dimensions. Section II of this chapter will discuss the orthodox vs subaltern divide, but I also want to note that deliberative model can also fall somewhere along a range between “contrived” and “diffuse”. I call deliberation "contrived" when citizens deliberate in artificial settings – citizen juries, town meetings, etc. I call deliberation "diffuse" when the it takes place in naturalistic settings within civil society. This distinction matters because each will have different requirements for deliberation to be legitimating, a point I hope to make clear in my chapters on etiquette and apology.

Contrived deliberation includes devices such as deliberative polls and citizen juries. In each case a group of randomly selected citizens meet for a few days in order to discuss important policy matters. In the former case the exercise purports to show how the citizenry might react if it had the opportunity to discuss an issue in-depth; in the latter the deliberators are expected to provide an advisory judgment to the government on a specific issue of interest.

An extreme version of this is the “Deliberation Day” proposal made by Bruce

Ackerman and James Fishkin. Deliberation Day would be an all-day event divided in four parts. First, participants would meet in neighborhood schools or community centers early in the day, where they would randomly be assigned into groups of 15. These groups would watch a live debate between the candidates on 2-4 important issues (selected by the candidates themselves). Next, each member would have an opportunity to speak and propose a question for a larger assembly. At noon, the group convenes with other small groups into an assembly of 500 people. A moderator collects the most popular questions from each group, curates them, and presents them to party representatives. At 2:00pm the representatives answer the questions and provide a brief recap. The citizens then reconvene into their small groups at 4:00pm and, again, each have 5 minutes to express their reactions to the answers.

Now contrast this structured, even fussy approach to deliberation with the vision that Robert Paul Wolff has of a democracy that managed to engage its citizenry:

America would see an immediate and invigorating rise in interest in politics. It would hardly be necessary to launch expensive and frustrating campaigns to get out the vote. Politics would be on the lips of every man, woman, and child, day after day. As interest rose, a demand would be created for more and better sources of news. Even under the present system, in which very few Americans have any sense of participation in politics, news is so popular that quarter-hour programs are expanded to half an hour, and news specials preempt prime television time.

Here there is no talk of formal meetings, contrived five minute speeches, and formal

---

352 Ackerman and Fishkin, “Deliberation Day.”
353 These are Ackerman and Fishkin’s numbers. I do not know why this larger assembly is not a multiple of 15.
355 Wolff, Defense of Anarchism, 36-37.
Q&A with party representatives. In Wolff’s world, deliberation is flexible and usually done with acquaintances in a relatively unstructured way.

This brings to mind part of what Habermas proposes in *Between Facts and Norms*. He divides public discourse into two tracks. One represents the formal institutional structures that translate public will into actionable policy. The other – and the one of interest here – is a diffuse discourse taking place within civil society. This discourse is not a Rousseauian image of a polity coming together. Rather, it is a case of institutions within civil-society that are more or less informal and that transmit the opinions through a series of filtering “sluices” toward the formal political decision-making institutions. Habermas believes that when this model functions successfully, democracies can avoid surrendering public will formation and implementation to technocratic interests while providing legitimacy in a modern context. For Habermas, legitimacy in modern democracies can only obtain when sectors of the public produce an inchoate set of voices which then get filtered and channeled into a form that democratic institutions can “hear”. Habermas’s model is in many ways the opposite of Deliberation Day, where the discourse is formal, contrived, immediately understandable by public bodies, and takes place between strangers.

The contrived vs. diffuse dimension of deliberation is a spectrum, not a dichotomy. Many proposals fall somewhere between the two poles. Some have argued in favor of public funding for civil society associations which encourage the type of deliberation that we collectively think is important. The idea here is very much in the

---

356 Habermas, *Between Facts and Norms*.
357 Ibid., 354-57.
spirit of Mill, where the idea is to promote citizens’ abilities to discuss matters intelligently, with an emphasis on reason-giving. Of course, there is no free lunch: what groups or issues count as “important” is itself a contentious decision that presupposes answers to questions of how deliberation should proceed. More on that in the next section.

At least one point, however, remains the same in all cases. For deliberation to legitimize, the discussants must demonstrate mutual (and self) respect. What this will actually entail in terms of concrete behaviors, however, will depend on the specific context where deliberation takes place. Arguing in a room full of strangers is not the same as navigating the politics of longstanding friendship groups, and in each of these cases the particulars matter. Discussing matters with some religious congregations may require more or less deference to the religious leader, and this itself depends on whether or not she is present.

I aim to discuss these matters more fully in the next two chapters, but first I want to address what may be the most important dimension of deliberation for my purposes: whether the deliberation should follow a priori rules, or the participants should be the ones to determine how the deliberation takes place during the process itself. The remainder of this chapter takes a closer look at how deliberative democrats have specified the rules of legitimizing deliberation. In doing so, I draw inspiration from André Bächtiger et al. by dividing deliberative theories into two camps. These camps, which I call orthodox and subaltern (and which the paper calls “Type I” and

---

358 Cohen and Rogers, “Secondary Associations.”
359 Bächtiger et al., “Disentangling Diversity.”
“Type II”) theories of deliberative democracy, form a continuum. To the extent that a theory is orthodox, it specifies a priori rules that can be used to judge deliberative procedures. To the extent that it is subaltern, it rejects establishing potentially exclusionary a priori rules and insists that the participants must themselves decide on the criteria used to judge the success of deliberation. In what follows I will discuss examples of each; in no way do I intend my account to be comprehensive.

II. Deliberative Models: Orthodox vs. Subaltern

II.1 Orthodox Models

Orthodox deliberative democrats are those which specify an ideal deliberative procedure and who use this idealized model to judge the quality of actual instances (or lack thereof) of deliberation. This does not mean that they insist that we necessarily ought to prefer a model that closely approximates the form of the ideal over one that does not. Any serious deliberative democrat must take into account the “problem of second best” – the notion that failure to achieve an ideal due to an inability to actualize one variable may require us to change other aspects of the decision procedure, even if these other aspects conform with what the ideal theory requires.\footnote{360 For the seminal paper on this topic, see Lipsey and Lancaster, “General Theory.”} To use a classic example: a failure to adequately address racial discrimination can turn the application of otherwise neutral rules into an even greater source of oppression. In such a case, justice may require us to turn away from desirable "neutral" rules and instead apply non-impartial rules which better serve the needs of justice in a particular situation.\footnote{361 For a forceful argument along these lines, see Marcuse, “Repressive Tolerance.”}
That said, theorists like Joshua Cohen have nonetheless put forth deliberative models to be used as templates for judging actual or proposed deliberative procedures, and I will focus on his seminal work as a model of what I call “orthodox” deliberative democracy. Since I will spend a good amount of time critiquing this type of deliberative democracy (mainly by way of the proponents of the “subaltern” family of deliberation), I will also spend some of time explicating it.

Cohen’s conception of an ideal deliberative democratic procedure consists of five features. He states that in such a model, deliberative democracy (1) is ongoing and assumed to persist into an unspecified future, (2) involves participants who are committed to making political decisions “within institutions that make deliberation possible”, (3) takes place within a pluralistic society, (4) makes clear the connections between deliberation and policy outcomes, and (5) involves participants who recognize each other as having deliberative capabilities, “i.e., the capacities required for entering into a public exchange of reasons and for acting on the result of such public reasoning.”

In addition to this, Cohen’s ideal deliberative procedure is characterized by four values. First, it is free in the sense that participants may consider the proposals on their own merits. What's more, the participants believe that the fact that a decision has been made via the deliberative procedure is reason enough to comply with it, regardless of what independent sources of authority may say or think. Second, deliberation is reasoned. Participants are to state reasons for advancing and defending

363 Ibid., 74.
proposals and expect that the proposals will be accepted or rejected on these bases. Unlike someone like *Theory of Justice* Rawls, Cohen – alongside the likes of Habermas – believes that in deliberative democracy “... collective choices should be *made in a deliberative way*, and not only that those choices should have a desirable fit with the preferences of citizens.”

Philosophers who cloister themselves in dark rooms to think through problems until a solution emerges need not apply – we need to actually get out and talk to real people in order to obtain deliberative legitimacy.

Third, Cohen's ideal deliberative procedure emphasizes *equality*. Each person is equally entitled to be heard, to put issues on the agenda, and so on. What's more, “The participants are substantively equal in that the existing distribution of power and resources does not shape their chances to contribute to deliberation, nor does that distribution play an authoritative role in their deliberation.” In other words, just as how in conventional democracy my vote is equal to the vote to my boss’s vote, when deliberating I should feel entitled to speak up without feeling cowed by their authority. By the same token, they should not pull their punches in a misguided attempt to encourage me.

Finally, ideal deliberative procedures aim for a “rationally motivated *consensus*” by finding reasons that are persuasive to all good-faith participants. Two observations should be made here. First, Cohen is aware that even in ideal conditions, this consensus might not be forthcoming. In these cases, the participants would resort

---

364 Ibid.
365 Ibid.
366 Ibid., 75.
to a vote. Cohen nonetheless believes that this type of vote is qualitatively different than that of a democratic system that merely tries to aggregate voter preferences. This is important for a variety of reasons, not the least being that, as Gutmann and Thompson argue, ongoing deliberation may be the best way to overcome misguided policy decisions.\footnote{367 Gutmann and Thompson, \textit{Democracy and Disagreement}, 77-79.}

Second, it is not clear if the persuasive reasons must be convincing in the same way for each participant. For instance, there may be many justifications offered for a policy proposal such as decreasing the tax burden on the wealthy. Someone may vote for it because it increases economic efficiency, another may do so because she believes that the level of taxation constitutes a violation of natural rights, and still another may support the policy because she is an elitist.\footnote{368 When not participating in deliberative fora, that is. Recall that Cohen wants us to express these values in the context of deliberative procedures.} The participants must be motivated by a desire to instantiate a common good, but there may be disagreement as to what that common good consists in.

Models such as Cohen’s have many attractive features. The most important salient trait of orthodox deliberation is that it specifies ahead of time what will prove to be stringent criteria for “correct” deliberation. There is a right and a wrong (or a legitimizing or non-legitimizing) way of deliberating, and deviations from those standards, while inevitable to some degree, are to be minimized and pointed out as flaws. Appeals that do not obviously appeal to a universally accessible sense of reason are out of bounds. This has the twin benefits of allowing for clearly delineated rules of
deliberation while protecting more vulnerable members from being the victims of capricious decisions cloaked in appeals to emotion or tradition.

What’s more, proponents of this type of model can argue that by only taking a vote when consensus is unachievable, participants are demonstrating respect towards one another. Democratic systems without deliberation may show at least some level of respect by giving each participant the same number of votes but do not combat the alienation that some voters may feel. In an aggregative democracy, an individual may feel that others are voting without considering her interests. Deliberation ensures that the disaffected have the opportunity to make their voices heard in a setting where they can feel confident that they will be taken seriously. By postponing the vote until after deliberation, all those involved can voice their frustration and demand direct, respectful answers from those who would maintain policies that continue to marginalize some participants.

Finally, in spite of their insistence on the need for a particular type of argumentation, deliberative democrats allow participants to agree for different reasons as long as those reasons adhere to the kind of deliberation participants have committed to undertake in the first place. As a result, the story goes, all earnest participants should have good reason to believe that they are being respected both qua rational actors who can participate in a reasoned discussion as well as holders of deeply held beliefs. By not insisting that everyone agree for the same underlying reason, deliberative democrats effectively bypass complaints that some individuals would be told that their deeply held perspectives are wrong or inappropriate.
II.2 Critiques of Orthodox Models and Rise of Subaltern Deliberative Democracy

Models such as Cohen's have come under sustained attack from a number of directions. Claus Offe, for instance, has pointed out that Cohen's model would be extremely hard to put into practice because it would make tremendous demands on would-be deliberators. In spite of their divergent aims, they are expected to “develop certain commitments and attitudes and must then act in accordance with these mental routines of preference formation.”\(^{369}\) It also means that individual deliberators must commit to the process with the knowledge that they may prove to be suckers if others duplicitously disguise their preferences.\(^{370}\)

But the principal attack against orthodox deliberative democracy have come from those who argue that its insistence on reason and "the force of the better argument" silence and/or exclude marginal groups. If these critics are right, this exclusion occurs because proponents of orthodox deliberative democracy define rationality in ways that invalidate the opinions of those whose communicative style and content, while natural to them, deviate from the proposed norm of deliberation.

This silencing need not be deliberate. As James Tully argues, dominant groups come to understand their position as being “uniquely reasonable”.\(^{371}\) As a result, members of privileged groups really do believe that they are being objective when they speak, failing to appreciate the importance of the cultural context that makes their

\(^{369}\) Offe, “Democratic Theory,” 100.
\(^{370}\) Ibid., 101.
\(^{371}\) Held, Models of Democracy, 243.
speech appear so natural, at least to them. But, as I will emphasize at length in the next chapter, deliberation must occur within a cultural context that leaves certain groups at a disadvantage.

One of the most important (if flawed) seminal works in this line of critique is Lynn Sanders' 1997 article, *Against Deliberation*.\(^{372}\) There she argues, among other things, that the models of deliberative democracy of her time did not take into account that “Some citizens are better than others at articulating their arguments in rational, reasonable terms.”\(^{373}\) What's more, she argues, deliberation also requires what she calls "epistemological authority," or "the capacity to evoke acknowledgment of one's arguments."\(^{374}\) The idea isn't merely that some people are more alienated or more unable to deliberate due to a lack of resources or past political exclusion, although any serious proponent of deliberative democracy would need to take these problems into account (and most do). Sanders is concerned with insidious prejudices that cause us to discount or “tune out” the opinions of some while perking up when others speak. Any of us can empathize with this line of argument at least to some extent by recalling a time when we noticed that the very polite face listening to us speak had glazed eyes and a frozen, distant smile.

Sanders is making two points as far as my discussion is concerned: (1) that people have different deliberative capacities, and (2) that some people's opinions will be discounted even when they perform their deliberative roles according to whatever

---

\(^{372}\) Sanders, “Against Deliberation.”
\(^{373}\) Ibid., 348.
\(^{374}\) Ibid., 349.
standard the deliberative procedure demands. Outward approximations of ideal deliberative procedures can hide the fact that the procedure is actually failing. Worse, since the deliberation seems to conform to Cohen’s criteria, it gives the outcome a veneer of legitimacy that would be absent had it been the result of naked coercion.

These points are expanded more fully in Iris Marion Young's work on deliberation and difference. Young's fundamental point, which extends across multiple works, is that orthodox deliberative democracy is flawed because it assumes we can transcend our particularities and come close to an objective, universal, rational perspective when deliberating.375 She argues that this assumption fails, however, for at least three reasons. First, it itself requires the heroic assumptions that people either (a) have not internalized beliefs about when they do and do not have the right to speak, as well as how to speak, or (b) can somehow slough off these beliefs when entering deliberative forums.376

Second, if we accept that people have different internalized beliefs about when and how they can speak, it becomes clear that the ostensible neutral ideal deliberative settings, even in their “pure” form, privilege some speakers over others. Specifically, Young argues that orthodox deliberative schemes, modeled after institutions such as formal scientific discourse, parliaments, and courts, reward agonistic modes of argumentation. As she says, “Speech that is assertive and confrontational is here more valued than speech that is tentative, exploratory, or conciliatory. In most actual

375 Held, *Models of Democracy*, 244.
376 Young, “Communication and the Other,” 122.
situations of discussion, this privileges male speaking styles over female.” I should point out here that Young is talking about communicative style, not necessarily communicative purpose. It would do no good for orthodox deliberative democrats to claim that the goal – consensus – is the antithesis of agonism if the method of achieving this consensus excludes swaths of potential contributors.

Young’s third and perhaps deepest critique of the idea that we can transcend our particularities rests in her skepticism that we can transcend group identification. Young’s earlier point about “male speaking styles” does not point to a belief in gender essentialism; in fact, Young explicitly rejects the validity of group identification via essentialism. She is instead arguing that our assigned group affiliation causes others to treat us in particular ways. This treatment, she argues, results in members of a given salient group sharing a perspective with other members of that group. For instance, whether or not someone identifies as female will not in itself cause others to classify them as such if that person is socially understood to be female. Others’ behavior towards that person – including threats, patterns of deference, and expectations which those who have not been assigned “female” will not face – will inevitably shape that person’s perspective. That person will be better positioned to understand the concerns of other people assigned to the category "female" than well-meaning "non-females" who have not been subjected to those same patterns of behavior by others. To put it simply: as a cisgendered male I can try my hardest to empathize, but I should not

---

377 Ibid., 123.
379 Ibid., 394-395, 399. I am frequently made aware of this point when my partner reports the numerous catcalls and aggressive body language displays that she faces on a near-daily basis.
delude myself into believing that I can adopt a perspective that contains the nuance of my cisgendered female friends’ lived experience. I certainly cannot do so in an ivory tower, without actually listening to them on their own terms.

The main upshot of Young’s critique is that our belonging to social groups means that our perspectives are necessarily partial. What’s more, Young believes that it is through actual deliberation or other good-faith interaction with others that we come to appreciate just how partial our perspectives really are. Deliberative democrats can neither enjoy the benefits of diverse perspectives nor show all participants equal respect if they insist on overly rigid rules of argumentation. Making demands that others speak “my language” won’t foster helpful communication. At best, it will provoke confrontation. At worst one of us will just shut down.

To make deliberation truly legitimating, Young suggests that the deliberative process be loosened by allowing three additional modes of communication. The first is greeting, i.e., statements that do not have propositional content, usually precede or initiate discussion, and demonstrate respect. This is a nod towards etiquette, which will be the subject of my next chapter. Young correctly points out that “Communicative interaction in which participants aim at reaching understanding is often peppered with gestures of politeness and deference, the absence of which is felt as coldness, indifference, insult.” As we shall see in the next section and chapter, this is not so straightforward, but she is right that any sensible ideal of deliberation must take these small courtesies seriously.

380 Young, “Communication and the Other,” 129.
Young’s second proposed mode of communication is *rhetoric*. Against deliberative democrats who argue that rational discourse – that is, discourse based entirely on making assertions and coolly giving reasons – should be the only permissible speech in deliberative proceedings, Young argues that rhetoric is necessary for some speakers to be heard, as it allows them to tailor the communication to their audience and prevent the onset of boredom. She believes that mere truth does not get always get the attention of the listeners. Sometimes you need to use humor, anger, drama, or wit to get the point across, especially when the audience is already predisposed not to take you seriously.

Finally, Young advocates the use of *storytelling*. The purpose of storytelling in this context is to help bridge the gap between interlocutors whose partial perspectives are causing conflict and who cannot understand the other person’s experience well enough to engage them productively. According to Young, storytelling allows individuals to share particular experiences that would allow others to better understand them. She furthermore argues that storytelling can help clarify important values, enabling groups to come to understand each other’s premises rather than beginning an acrimonious discussion on the mistaken assumption that the groups share the same fundamental goals. What’s more, she believes that storytelling allows everyone to see the reaction others have to particular narratives, with the idea that the sharing of a multiplicity of narratives and the reactions to them can form a “collective social

---

381 Ibid., 130.
382 Ibid., 131.
383 Ibid., 131-32.
wisdom not available from any one position.”

In sum, Offe, Sanders, and Young are among those who I have classified as “subaltern” deliberative democrats. They reject what they take to be overly rigid rules of deliberation proposed by orthodox deliberative democrats, arguing that such rules systematically further marginalize already disadvantaged groups. I am sympathetic to their project, yet, as I will argue in the following section, I believe that their work is also subject to strong critiques. The model that I will ultimately propose is in the spirit of the subaltern deliberative democrats, but with an emphasis on the importance of private assessments of the legitimacy of the proceedings.

III. Critiques of Subaltern Models and Impasse

While subaltern deliberative democrats should be lauded for pointing out real problems with the orthodox models, they nonetheless fail to fulfill the aim (as I understand it) of deliberative democracy, at least as far as this dissertation is concerned. Deliberative democrats want – or should want – the process of deliberation to convince participants to respect the outcome of the decision procedure. Subaltern deliberative democrats believe that this can happen if deliberative models are inclusive. By allowing all kinds of speech (that are at least not blatantly disrespectful), they hope that all would be incentivized to participate in the deliberation and to come out the other end feeling as though their input was meaningful.

I agree with this intuition, with the addition that a commitment to inclusiveness

---

384 Ibid., 132.
would activate would-be deliberators’ duty to deliberate. After all, if there is a prima facie duty to deliberate (as I have argued in the previous chapter), this obligation would come into question if someone were to have good reason to believe that "their kind" would not be taken seriously in the deliberative forum. If deliberators do not have good reason to believe that they are being excluded from the start, however, then a decision to abstain from participation would constitute a failure to uphold one's duty and would severely weaken the argument that the results of the deliberation do not apply to the abstainer. So far so good, but I argue here that the subaltern commitment to inclusiveness either robs deliberators of the opportunity to come to any agreement or it merely kicks the inclusiveness can down the road.

Consider first the prototypical – even stereotypical – ideal orthodox deliberator interacting with the type of deliberator that subaltern theorists seem to have in mind. The former is committed to the ideals of deliberative democracy and is able to accept that others do not speak like him, i.e.: like a philosophy professor. By training and temperament, he speaks in terms of and is persuaded by what he considers to be rational arguments that "all should be able to accept," given enough time and reflection. His interlocutor responds to his arguments with passionate but meandering narratives. Try as he might, the first speaker cannot see the point in the narrative and is unsure how exactly to respond. The second speaker is having just as much trouble; not being used to the first speaker's style, he responds with more speech which would

385 It might be helpful to think of this erudite philosopher, when visiting his cousins during the holidays, posting exasperated Facebook updates complaining about just how hard it is to communicate with these well-meaning people who have had very different life experiences.
be readily understood by his own peer group.

What does the first speaker do with these authentic expressions? Two immediate possibilities suggest themselves. The first is to say that this is silly and that he can understand the other person without much difficulty. We can call this the (exasperated) “Come on, we get along just fine in the real world” objection. But if this is true, then the objections raised by subaltern deliberative democrats should not be seen as having much weight. To the extent that the two interlocutors can easily bridge the gap with a narrative or two, the subaltern objection essentially turns into a request for a bit more time during deliberation and a recognition that narratives, storytelling, and the like can contribute valuable information – easily understood by all – that would not be available under more restricted speaking rules. This line of thinking essentially relegates the subaltern project to a footnote. I disagree with this move; the problems that subaltern deliberative democrats raise are pressing enough to warrant more than an eye-roll. In fact, the real problem is that subaltern solutions don’t go far enough.

“Agreement” can mask the fact that one party seems to always compromise more than the other, a topic that I will address at greater length in the next two chapters. Countless people have seen their marriages end suddenly, oblivious that their soon-to-be-ex deeply resented the arrangement that seemed to be working so well. Countless others have quietly endured humiliating conditions, their objections (when made) totally misunderstood or discounted by a spouse who sincerely thought that egalitarian communication and compromise was taking place.
Yet subaltern solutions don’t always fix the problem. John Dryzek gives a number of examples where subaltern "expansions" to deliberative democracy, notably those made by Young, can lead to failure. Storytelling itself is subject to group norms that evaluate what counts as "correct" and "incorrect" storytelling. There is a danger that some authentic attempts at storytelling "cannot easily withstand the normalizing gaze of the group." Greetings can be coercive (in the case of, for instance, an overly strong handshake, complete with eyeballing) or exclusionary (for example, in the case of an exclusive handshake or gang sign). Finally, deliberative democrats have long worried about the destructive potential of rhetoric, and while excellent rhetoric can unite those with disparate viewpoints, it more often than not "requires a context of a community of like-minded people, taking force only ‘when people are united in their aims', igniting their passions and making it less likely that their differences with those outside the group can be resolved." We might try to say that deliberators can solve these problems with more storytelling, greetings, and rhetoric, but here we have no guarantee that these efforts won’t contribute even more to an oppressive environment. When modes of communications are the problem, talking it out can make the problem worse.

The second way of approaching the philosophy professor’s problem is to say that the speakers can understand each other, but only with much effort. Again, this solution is problematic. How do the speakers eventually reconcile their seemingly

---

386 Dryzek, Deliberative Democracy," 68.
387 Ibid., 69.
388 Ibid., 70.
incommensurate modes of speech? If they are never able to do so, then deliberation has clearly failed. The philosopher philosophizes, the interlocutor tells a story, and they both shrug and say "Now what?" Orthodox deliberation, for all its faults, at least specifies a method that can adjudicate between different types of speech. Of course, this is the kind of ostensibly a priori adjudication of what counts as "proper deliberation" that evoked subaltern objections to begin with. Subaltern deliberative democrats – to the extent that they are subaltern deliberative democrats and not merely proponents of orthodox deliberation with a few bells and whistles – cannot propose a standard by which participants determine what to do with, for instance, "greetings".

The upshot is that unless the participants can make themselves understood to each other, they will be at a standstill. They could resolve this by resorting to a vote, but then deliberation would lose its decision-making dimension. Each party can honestly say that they have been given the opportunity to express themselves, but this expression neither leads to an outcome nor is directly connected to whatever outcome emerges. I say my piece, you say yours, we mostly talk past each other, and we vote. Why not just skip the preliminaries and get on with the vote?

Once again, at least two responses are possible here. The first is that there are epistemic gains stemming from deliberation even if the participants never come to an agreement or even fully understand each other. Fair enough; increased earnest

---

I earlier suggested that preceding voting with deliberation might constitute a sufficient show of respect so as to bind all participants to the result of the vote. This is not what I am describing here. Here I describe cases when the vote follows what essentially amounts to a failure to communicate effectively. In these cases, those on the wrong side of the vote could justifiably claim that they were not "heard," and that the deliberation was therefore fundamentally ineffective.
dialogue probably gives the parties at least some information regarding the others' preferences. In other words, the philosopher may never fully (or even largely) understand the interlocutor, but neither does he view her as a gibbering Martian. Some progress is made.

Yet this benefit is simply not enough to make the deliberation worthwhile from a legitimating perspective. If deliberation is *infrequent*, as is likely the case in contrived settings, then it is unlikely that enough dialogue will take place for marginalized participants to "feel heard," especially if the miscommunication stems in large part from different modes of expression. The ten cumulative minutes of floor time that participants in Deliberation Day get every major election is clearly insufficient.

Deliberative democrats who are mainly concerned with the epistemic value of deliberation may be satisfied since the resultant policies will be more likely to be correct\(^\text{390}\), but these are not the types of arguments that would likely sway philosophical anarchists. If on the other hand deliberation is frequent and lengthy but rarely "gets it right", we are now asking marginalized citizens to spend lots of time speaking with people who, through no fault of their own, frequently misunderstand

\(^{390}\text{Assuming, that is, that "correct" either means (1) "better conforming to the actual preferences of the participants" or (2) "in line with some sort of objectively correct outcome," with the proviso that more communication is likely to get us closer to that outcome. Those who espouse this second interpretation should take into account the relative weight of the resultant policy outcomes. For example, maybe deliberation raises the odds of coming to a correct decision on one class of decisions from 55% to 65% but decreases the odds on another class (perhaps those that are especially prone to rhetoric) from 50% to 45%. Even if the first subgroup is larger than the second, the importance of the second may far outweigh the first. Thus, proponents of deliberation on these groups would need to show or at least argue that deliberation is either a Pareto improvement or that the weighted good of deliberation outweighs the weighted harms it brings about.}
(often in insulting ways) the marginalized citizens’ speech.

The second possible response is that there is expressive value in deliberation, that deliberators get something out of having the opportunity to speak their mind, even if others misunderstand. I am sympathetic to this perspective and will draw on it to some extent when I propose an alternative model of deliberation in the next section and which I examine in subsequent chapters. But for now, it is hard to see how the opportunity to express one's interests in one's own terms would lead to legitimation if one is consistently misunderstood or otherwise marginalized, inadvertently or not.

In sum, the problem with subaltern deliberation as I understand it is that it leads down one of two paths. One the one hand, if it relies on an implicit mutual understanding that spontaneously occurs as a result of an expansive discussion, it runs the risk of recreating the orthodox deliberation’s problems. Things "somehow work themselves out," except that some members inevitably exercise more rhetorical or social power than others. Worse yet, since the standards have not been articulated in advance, abuses of power are even more difficult to "call out" than with a standard orthodox model.

On the other hand, if the procedure strives for purity of inclusion and celebrates the presence of diverse modes of communication without making an effort to bridge the gap to make the communicative styles mutual intelligible, it can no longer lead to actual decisions. In the extreme case, the participants resemble Baugh's Wolffians, which I discuss in the final chapter, where everyone solipsistically makes their communicative performance and no meaningful engagement occurs.
That said, it seems clear that claims of incommensurability could go too far. In the final section of this chapter I outline a sketch of deliberative legitimation that I believe takes the problem of intelligibility seriously without going so far as suggesting that the participants have somehow lost the ability to make themselves understood on a basic level. In essence, I will suggest that we accept that misunderstandings – very important misunderstandings – may take place and that we might not know how and when they do. I also suggest that this does not matter. What counts, I argue, is that all participants feel respected even if they suspect that they do not "really" understand each others' perspective. Since participants have different backgrounds and beliefs about what constitutes respect, each will determine for him or herself the standards which must be met for deliberation to be legitimizing to her. In order to present this model, it would be helpful to begin with a discussion of privacy in the context of deliberation.

IV. Privacy and Modified Model

IV.1 Privacy and the Internal World

Deliberative models of all stripes are interested in reciprocity. As Gutmann & Thompson say, “Reciprocity is widely recognized as a core principle of democracy in its many moral variations – liberal, constitutional, procedural, and deliberative – but most theories do not give it the central role that deliberative democracy does. Reciprocity holds that citizens owe one another justifications for the mutually binding
laws and public policies they collectively enact.”391 This "right to a justification" differentiates deliberative democracy from its competitors, but it does not, so far as I can see, adequately explain why people are owed such justifications. Surely it has something to do with respect, but this supposed “right” is itself problematic. Indeed, it runs afoul of the commonsensical notion that I have the right not to justify myself in many circumstances. Deliberative democracy – especially subaltern deliberative democracy – seems to have a disclosure requirement that gets it into trouble.

This disclosure requirement at times runs deep and is instructive. Benhabib's mostly subaltern account relies heavily upon it. She argues that when confronted with moral disagreement, we should articulate our needs. In doing so, we are to determine which of those needs are actually false, a process that requires a deep plumbing of one's psyche. Donald Moon describes her position as follows: “By discovering that one's position is rooted in 'the dark recesses of the psyche,' one will be in a position to reevaluate one's claims as a result of this deeper self-understanding. This, moral-transformative discourse leads to a kind of self-transparency which can form the basis of a genuine community.”392 Deliberation, especially in subaltern accounts, moves past a respectful articulation of needs and comes close to self-discovery done in a group context.

Moon is rightly disturbed by such approaches. He points out that first, such a deep excavation of meaning implies that these meanings are somehow “out there” and can be apprehended. But this is dubious. On the contrary, Moon correctly points out

391 Gutmann and Thompson, Deliberative Democracy, 98.
392 Moon, Constructing Community, 89.
that the very act of apprehending and communicating those needs transforms them.

“Our motivations and the meanings of what we say and do are not univocal, fixed
things, but exist on a variety of levels, in complex relationships to various aspects of
the self, and are fraught with ambiguity. In an important sense, then, we can never
fully articulate our needs.” Deliberation ceases to be about clear articulation and
becomes an exercise in preference formation – precisely the kind of activity that
frightens philosophical anarchists or indeed anyone who wants to maintain trust in
their independently determined judgments. This is not a case of merely being
persuaded by a well-reasoned argument, but one where the stability of the
philosophical anarchist’s conscience is undermined.

Given these considerations, the subaltern deliberative democracy project
appears confused. On the one hand, perhaps they are trying to have it both ways:
transformation while respecting the integrity of individual experiences and
perspectives of the (good-faith) participants. Or, on the other hand, perhaps they
believe that needs and preferences can exist independently of their expression. On this
line of thinking, all we need is the right kind of honesty, willpower or therapeutic
techniques to bring them to the fore. But this simply ignores the ostensibly therapeutic
and clarifying nature of deliberation that Benhabib describes.

Moon's second objection is that unconstrained discourse is itself potentially
coercive. Coming to terms with one's “real” wants and needs, even if this is fully
possible, is usually a very demanding task even in the most supportive environments.
In cases where the “probing” of one's wants is done in less supportive settings, “The
implicit sanction of the disapproval of others, pressures to conform, and the fear of revealing one's weaknesses, or leaving oneself vulnerable in other ways, may make it necessary for individuals to withhold or even dissimulate their beliefs and feelings, and often can prevent them from recognizing their own needs.\textsuperscript{393} In other words, when put on the spot, we tend to be unable to recognize and articulate our wants and needs, and even if we could, we would be in precisely the position that would cause us to want to hide our desires. In order to fully come to terms with ourselves, we need privacy, not an audience.

What’s more, Moon argues that privacy is necessary for autonomy, friendship, and even the creation of a coherent identity, precisely the traits that deliberative democrats want to foster. This is especially true when we live in a variegated society where we are expected to play multiple roles across various spheres. In such a context, he argues, “to maintain a sense of self and a more or less coherent set of values and principles requires that we be able to control the ways in which we reveal ourselves to others.”\textsuperscript{394} Not only is this important in order for us to develop and maintain a coherent psyche, but the selective divulgence of personal information is part of what makes friendships so valuable. Moon points out that revealing ourselves to friends is a “way of showing one’s commitment to another person and of creating and sustaining a deep relationship with him or her. If one seeks to have this kind of intimacy in some relationships, certain forms of self-revelation will be inappropriate in others.”\textsuperscript{395} Those

\begin{itemize}
\item \textsuperscript{393} Ibid., 92.
\item \textsuperscript{394} Ibid.
\item \textsuperscript{395} Ibid.
\end{itemize}
deliberative democrats who push for more self-disclosure must not only show how it would enhance the deliberative project; they must show how it would do so without undermining the development of a coherent identity, mature views, and the capacity to hold some information as “special”, to be shared only with those close to us.

Proper deliberation involves reflection, and this reflection must allow for (1) the withholding of information and (2) the right to not generate an opinion – even to one's self – when doing so would interfere with the process of self-reflection. To use an analogy: creating an identity or coming to certain types of opinions is like baking a cake. Continually opening the oven door to check on the cake or turning up the heat to make the cake bake faster do not expedite the process – they increase the chances of ending up with a deformed or poorly cooked dessert.

This need for privacy, even (or possibly especially) within deliberative contexts, provides an entry point for a reconceptualization of deliberative democracy based on the philosophical anarchist’s individualist framework. For my purposes, what is crucial about this right to privacy is that if we grant (1) that we have a duty to deliberate and (2) that we have a right to privacy, it does not follow that we have a right to make binding moral judgments on whether or not a given individual is exercising her right to privacy appropriately. Doing so would require her to disclose the contents of metaphorical “oven” and would thus risk ruining the cake. Failing to take privacy seriously, including respecting individuals’ right to make (and not necessarily disclose) private judgments on the proceedings would undermine the deliberative project.
What’s more, maintaining control over access to one’s “internal world” is intrinsically valuable, not a merely instrumentally valuable way to improve deliberation. If the duty to deliberate has led us to a duty to relinquish control over the integrity of our internal world, something has gone seriously wrong.

Yet it is worth bearing in mind that these concerns apply to subaltern approaches, not orthodox ones. The very formality that annoys the subaltern deliberative democrats prevents orthodox models from violating the deliberators’ inner sanctum. What is needed, then, is a model of deliberation which respects the sanctity of deliberators’ privacy and autonomy while allowing everyone to participate in a way that they understand to be meaningful.

IV.2 Recap and Philosophical Anarchist-friendly Model

It is now time to recapitulate some of the most important points. I suggested in the previous chapter that we may have a duty to deliberate, which led me to examine some of the literature on deliberative democracy in the hopes of finding a model that would force philosophical anarchists to acknowledge a duty to obey laws stemming from legitimizing deliberation. Some deliberative schemes were out of bounds from the start since they presupposed a legitimate political system. I therefore focused my attention on “orthodox” and “subaltern” models that seem to say that deliberation is a necessary component to legitimacy, and not the icing on an already legitimate cake.

Orthodox models propose deliberative procedures that emphasize rational argumentation, with participants expected to give reasons that could be convincing to
all. Subaltern critics respond by arguing that deliberative proceedings cannot occur in a social vacuum and that the rules of the proponents of orthodox deliberation are not so neutral after all. They instead propose models that they take to be more inclusive and more sensitive to the deliberators’ differences. This, they hope, will help deliberators avoid a false consensus, one based on implicit domination. Finally, I argued that while I am sympathetic with the subaltern deliberative democrats’ project, their proposals do not solve the problem of philosophical anarchism.

The fundamental problem with subaltern deliberative schemes lies in their reticence to abandon the hope of engaging in deep communication with others. The search for this deep deliberation risks leading them down one of two problematic paths. By advocating the inclusion of disparate modes of communication, they either (1) risk failing to find a standard by which to harmonize and adjudicate the different types of discourse, or (2) hand-wave and assume that the participants will work it out, and in doing so turn their model into a souped-up orthodox model, complete with external (and probably poorly specified) standards by which to judge the legitimation of the proceedings.

I believe that a model of deliberation that would satisfy philosophical anarchists should abandon the aim of having deliberators deeply understand each other. Some may believe that deep communication is necessary to demonstrate respect, but this need not be the case. What ultimately matters is that each deliberator feels respected, a feeling that does not march in lockstep with a belief that one is being understood.
Many people neither expect nor care about others’ understanding in some contexts. Philosophical anarchists could believe that their need for respectful treatment is satisfied even if they are being profoundly misunderstood and know it. Ultimately, each deliberator must decide for herself whether or not the deliberation is occurring respectfully and, if not, whether or not the disrespect means she can disengage (literally or figuratively) from the proceedings with the belief that she discharged her duty to attempt to deliberate respectfully with those around her.

This approach has a number of virtues. First, by re-emphasizing the individual, it respects the common atomistic philosophical anarchist framework. This means that no additional arguments are necessary to shift the philosophical anarchists’ perspective. Second, it still appreciates the types of group-based considerations brought up by the likes of Young. For example, it is possible that members of a social group feel unheard or disrespected by deliberative proceedings. It is possible that through discussion, most of these members come to feel respected by the other group members. Yet any given individual in the group may have a good reason to believe that the attempts at understanding made by other deliberators were insufficient to demonstrate proper respect. This person is entitled to this belief, as long as it is attached to a good-faith moral framework that acknowledges the dignity of others, and the fact that others in her group feel respected shouldn’t override her belief that deliberating would, in this case, violate her sense of self-respect.

As I have discussed in the previous chapter and will discuss at greater length in the next one, the fact of reasonable pluralism implies that people may come to have
different good-faith standards as to what type of behavior is necessary for them to believe that they are respected. Rather than let the proceedings go on interminably or announce that everyone should have hypothetically consented to the outcome because most deliberators think that the proceedings have taken place appropriately, it would be more honest to simply say that the output of the deliberation is illegitimate with respect to certain individuals. As I will discuss in the following chapters, the standards for what constitute respectful deliberation deal with not only the content of the deliberation, but the small, oftentimes nearly imperceptible actions that inform how real discussion actually takes place.

Before turning to those questions, I would like to end by flagging and addressing two challenges to this type of response to the philosophical anarchist. The first challenge is that even if deliberation does legitimize, the people or entities to whom we have political obligations will likely not track existing state boundaries. For instance, it is unclear why a would-be philosophical anarchist living in Vermont would have a greater duty to deliberate with a Texan than with a Quebecker.

I would be happy to concede this point. I see no reason why archists should insist that the current political boundaries are the only appropriate ones. If philosophical anarchists come to accept that they have political obligations to, say, local communities, progress will have been made. Acknowledgment of one’s political obligations to others does not imply acquiescence to a system of borders that purports to link disparate peoples into a state. This is not to say that people in large countries

---

396 Recall that in chapter 1 I drew on A.J. Simmons' idea, based on his interpretation of Locke, that the state can be legitimate with respect to some people while not being legitimate to others.
cannot feel political affiliation towards each other, but rather than there is no error in believing that one’s political allegiances do not track the existing state.

The second challenge is that this model concedes too much to dissenters. Vociferous racists have nakedly unreasonable criteria that others must meet for them to feel respected. A racist might say that only members of their race should be able to look them in the eye. A sexist might be offended if a woman speaks to him assertively. Why should these unreasonable people have fewer duties than those who hold morally decent views?

I argue, possibly counter-intuitively, that these people are in the same position as philosophical anarchists with respect to political obligations. Neither one group nor the other is part of the political community proper. In other words, neither of them has political obligations – a subset of moral obligations – towards the group. In both cases, the dissenters are being coerced, possibly justifiably so. The difference is that philosophical anarchists who publish academic papers do not advocate obviously immoral positions. As a result, the force of coercion is rarely felt in those cases – perhaps we could say that the state is latently coercing them.

This is not the case with racists who believe that the laws are illegitimate because of the presence and participation of racial minorities. They feel the force of coercive laws because to them the laws are not merely superfluous dictates that endorse what the racists already wanted. The racists are forced to live in a way they find objectionable.

I use “forced” very deliberately here. To say anything else – to try to say that
they “hypothetically consented” to a system they find repugnant – would be to disrespect their autonomy. It is possible for X to coerce Y while believing that he is respecting Y, even if Y disagrees. All X can do is implement moral rules as best he can, doing his best to show appropriate respect to all. To coerce Y and to deny the truth of Y’s assertions that she does not believe that she and X share a political community, however, goes too far. Racists, like philosophical anarchists, are entitled to the dignity of maintaining their sense of identity, including their sense of which groups shape this identity, even when being forced to live in accordance with the law.

To emphasize the point: exclusion from a political community does not imply exclusion from the moral one. As I explained in Chapter 5, our personhood sets a floor on how poorly we can treat others. Those who are excluded from our political community are still entitled to a minimal level of treatment, a level that might be quite high. Members of the political community cannot persecute those who are excluded, and those who are excluded are not only subject to justifiable coercion by the members of the community, but are also bound by moral law stemming from the existence of the dignity of others. Political exclusion does not give anyone a free pass to run roughshod over others.

On the other hand, this modified model does not imply that someone's sincere belief that she is being disrespected mean that she is not subject to collective decisions. What it means – what this dissertation ultimately argues – is that in such cases she is not bound qua member of the political community. To take it from her adversaries' perspective: if we (1) make good-faith efforts to engage others in
dialogue, (2) sincerely believe we are treating them with respect, and (3) believe that in certain cases (possibly all cases) it would be morally correct to coerce those individuals, then we have the moral right – and responsibility – to do so.

I now want to close this chapter by discussing some of these conclusions with reference to Thoreau’s *Civil Disobedience*, where he advances a position which has been interpreted by some as philosophical anarchism. While this claim is contested, it is understandable, given passages like: “I think that we should be men first, and subjects afterward. It is not desirable to cultivate a respect for the law, so much as for the right. The only obligation which I have a right to assume is to do at any time what I think right.” Thoreau seeks neither the total abolition of government, nor the existence of government on principle, but he does believe that the government in place at the time of the essay’s publication was sufficiently unjust to warrant resistance. He argues throughout his essay that conscience dictates that we have a duty to disengage from such a government, withdrawing support whenever possible. Yet Thoreau recognizes that the consequences of this disengagement from the state include not only material hardship, but social alienation as well. Upon his return from jail, he

397 Fiala, *Political Skepticism*, 90.
399 Ibid., para 7-8. Thoreau accepts the existence of a government that acts according to the majority’s opinion when it comes to matters where it appropriate to express one’s “feeble interest,” but does not think that majority opinion should matter when it comes to matters of right. See para 12.
400 Ibid, para 14, 19, 38. Thoreau is sensitive to the tensions between the desire to disengage and living a life that entails at least basic social connection. *Civil Disobedience* contains numerous exhortations to live simply and to resist the trap of wealth and status, yet Thoreau persists in wanting to be a good neighbor in spite of his disappointment and feelings of alienation when he returns from prison. And while he declares that he wants to only be a member of those organizations to which he consented to join, he also famously says that “If I had known how to name them, I should have signed off in detail from all the societies which I never signed on to; but I did not know where to find a complete list.” See para 27.
finds that his neighbors’ “was for summer weather only; that they did not greatly propose to do right; that they were a distinct race from me by their prejudices and superstitions, as the Chinamen and Malays are […]”\textsuperscript{401}

Thoreau thus appears to be the model philosophical anarchist: conscience-driven, suspicious of subordination to group norms, and unimpressed by collective decisions. He also appears to adhere to the philosophical anarchist precept that there is no obligation to dismantle a basically just government that restricts its activities to a sphere that does not trespass on the dictates of conscience.\textsuperscript{402} When one’s support of the state interferes with conscience, however, Thoreau asks that we decide how to act based on our own moral principles, even if this entails going to jail.

Yet Thoreau’s apparent philosophical anarchism is complicated by his sensitivity to the tensions between the desire to disengage while living a life that entails at least basic social connection. \textit{Civil Disobedience} contains numerous exhortations to live simply and to resist the trap of wealth and status, yet Thoreau persists in wanting to be a good neighbor in spite of his disappointment and feelings of alienation when he returns from prison. And while he declares that he wants to only be a member of those organization to which he consented to join, he also famously says that “If I had known how to name them, I should have signed off in detail from all the societies which I never signed on to; but I did not know where to find a complete list.”\textsuperscript{403}

\begin{footnotes}
\item 401 Ibid., para 35.
\item 402 Ibid, para 19-20.
\item 403 Ibid., para 27.
\end{footnotes}
Here Thoreau is recognizing his membership in a political community, but is trying to weaken the connection between this community and the state. When Thoreau wants to be a good neighbor, it is not merely out of a sense of cosmopolitanism, nor because of some independently derived imperfect duty of charity. He appears to want to be a good member of his community. The tension remains unresolved, however; his deep disappointment in his neighbors does create (or bring to light) a distance between the two. His best attempt at resolving the tension between affiliation and alienation is to posit the existence of a people, to claim that the state is connected to these people in only a superficial way, and then to claim that the people can be redeemed by those who help put an end to injustice by letting their lives “be a counter-friction to stop the machine.”

What I propose is that men of good conscience may find it necessary to jail Thoreau, and not because he is failing to fulfill his ostensible duties as a citizen, as a standard archist would have us believe. Thoreau’s feelings of alienation work both ways. He cannot reject a state-mediated connection with his neighbors and dictate the terms of their affiliation when they believe that the state plays a crucial part of their identity as a people. Put simply: many would disagree with Thoreau’s description of the fragile connection between the government and the identity of the people, and this should count for something. With that in mind, I believe, unlike those who would say either that (1) Thoreau is right, that the state is or should be irrelevant when describing a people, or (2) that Thoreau is mistakenly exaggerating his claims of alienation

404 Ibid., 19.
(notwithstanding his comment about “not being able to find” all the groups that he had joined), that Thoreau is alienated, and is trying to smuggle in a political identification with those who would consider him an outsider if they understood the depth of his alienation and rejection of the state. You don’t get to be part of the club if you’re flouting rules that the other club members feel are constitutive of membership.

To be clear, I am not saying that there is no affinity between Thoreau and his community, that he may as well be foreigner who has stepped onto American soil for the first time. In a sense, being an American (or whatever) is not a purely black-and-white proposition, legal designations notwithstanding. An American who moves abroad and changes citizenship does not suddenly stop being recognizable as an American, both to herself and others. Yet she is not American in an important sense. She’s off the team, so to speak, regardless of the level affinity she feels towards her former country. She no longer has the same duties and, reciprocally, Americans have the right to consider her an outsider, albeit a familiar one.

While Thoreau’s alienation (as I have described it) stems from the gap between his inability to support an evil regime in good conscience and the acquiescence of his neighbors, my focus is on the alienation that comes from or is perpetuated by disrespect by those with whom archists say we have a political relationship. To reiterate my main point: we have a duty to deliberate with others in order to adequately discharge our primary duties, but we cannot in good conscience participate in deliberative practices that undermine our self-respect. As with Thoreau, the question is a matter of conscience. The alienated person cannot be forced to believe that she is
being respected (although others could try to persuade her that she is wrong), and those who do believe she is being treated respectfully must decide how to treat alienated individuals in their midst. If they believe that Thoreau should be forced to abide by the rules that apply to the members of the political community, so be it. Absent a method that can convince one side or another to change their views, all Thoreau can do is disengage, and all the political community can do is determine how to act towards people like Thoreau. What they can’t do is deny the domestic alien’s insistence that her self-respect prevents her from being a full member of the political community.
Chapter 7: Etiquette

This chapter brings together some of the most important points of the previous two chapters. Earlier I suggested that natural duty could provide philosophical anarchists with the type of argument that could persuade them to become archists. I further proposed that philosophical anarchists might agree that they have a natural duty to try to deliberate in order to discharge their first-order duties. I then made a brief overview of the deliberation literature and argued that much of it either objectionable or does not apply to the question at hand – such as when it seeks to optimize a system that it presumes to be legitimate – or provides overly general guidance on how deliberation is to proceed.

Here I drill deeper by considering the role of etiquette, politeness, and related concepts in deliberative processes. There are two parts to the argument. First, I differentiate between polite behavior, impolite behavior, politic behavior, and face. Much of the existing literature, drawing on Brown and Levinson’s (B&L) canonical Politeness, has more or less equated polite behavior with face-saving behavior. I reject this view and instead side with Richard Watts’s critique of the B&L model, adopting his idea of "politic" behavior as the central concept describing ordinary interactions. Politic behavior, not necessarily politeness, is what deliberative democrats probably have in mind when imagining deliberation. That said, I will try to show that they have difficulty specifying when politic behavior should be maintained.

Second, I take laughter as an illustrative case to highlight tensions inherent in
“polite”, “impolite”, and “politic” behavior, showing how there is no a priori way of specifying what should count as morally desirable behavior in any normal case – i.e., in any case where the content of the action clearly represents a violation of duty. People develop rules about how they should be treated (and how they should treat others), and this bears heavily on whether or not they can accept with a clear conscience the social conventions that govern normal interactions.

I. Politeness, Face, and Politic Behavior

I.1 Brown & Levinson

I.1.a Basic theory

The most straightforward way to apply a theory of politeness to deliberative democracy would be to develop a theory and then to use it to generate maxims that constrain what counts as legitimizing deliberation. It would, in other words, disqualify some attempts at deliberation. Such a theory would have at least two requirements. First, it would have to be sufficiently general to apply cross-culturally. This means that a theory that generates maxims for deliberators in France would have to be able to generate maxims for deliberators in New Zealand, although local conditions may cause us to create different maxims in each case. Second, the theory would have to have sufficient bite to allows us to generate maxims that translate to specific behavioral imperatives ("Do this, don't do that") that can enable us to determine whether or not the maxims are being observed. "Be respectful" is too vague. "Do not use profanity" is much better. “Don’t use this list of words” is better still.
Although there are many competing theories of politeness, the towering modern work in the field is Brown and Levinson’s *Politeness*, and this is the book I will take as a model of the type of theory of politeness described above. Not only are many alternative theories of politeness derivatives that branch off from this work, but B&L incorporate previous seminal works, including Erving Goffman’s writings on face, without merely repeating them. As a well-respected, comprehensive theory, *Politeness* thus seems to offer us the most obvious hope of finding a model of politeness that we need to buttress the legitimating function of deliberation.

The central conceptual unit in B&L model of politeness is a Model Person (MP) who is instrumentally rational and who has a, and is concerned with, "face." This idea of "face" is central to B&L’s theory, and is defined in a preliminary way as “the want to be unimpeded and the want to be approved of in certain respects.” B&L believe that the MP has two faces. The desire to be able to act unimpeded – what political theorists might call negative liberty, is called "negative face." "Positive face" is a bit more ambiguous, but essentially consists of "the positive consistent self-image or ‘personality’ (crucially including the desire that this self-image be appreciated and approved of) claimed by interactants."

Since we all value our face and we need others’ cooperation to meet our face needs, the theory posits that we have good reason to help each other maintain face. B&L call any action that threatens the face of one of the interactants a *face-*

---

406 Ibid., 58. Their idea of face draws heavily on Goffman, *Interaction Ritual*.
threatening act (FTA). They argue that because of the need for cooperation in meeting face needs, we will typically perform *facework* to help others fulfill their face needs with the expectation that they will reciprocate throughout the interaction. It is a case of "I scratch your back, you scratch mine," or, more typically, "I won't make you look dumb if you don't make me look dumb." Unless someone is in a situation where she can trick or coerce someone to meet her face needs without the need for reciprocation, she will typically work hard to avoid FTAs or to perform reparative facework if a violation does occur.

That said, Brown and Levinson are well aware that FTAs obviously *do* occur, and not always accidentally; they are sometimes foreseeable results of one’s actions. Often we cannot meet our needs without risking damaging others' face. For example, when a student starts muttering loudly in class, a professor can clearly express her annoyance and signal the behavior she wants by telling the student to shut up. This, of course, is the opposite of face-saving behavior but has the virtue of being clear and (in some cases) more likely to elicit compliance.

There is thus a tension between the needs to achieve one’s prosaic aims and the need to help others save face. This amounts to an optimization problem. In this case, B&L would suggest that the professor will weigh what she is trying to accomplish by

---

408 Ibid., 65.
409 Ibid., 60.
410 Since I am ultimately dealing with legitimizing discourse, I am not speaking about statements whose sole purpose is to injure or humiliate. I assume that those remarks violate basic natural duties. For example: I am rejecting out of hand statements like "You're an idiot" if they are said maliciously or gratuitously. I am however not immediately rejecting the same statement if it is said in a context where the speaker believes that the predictably insulting comment conveys important propositional information.
the potential FTA against the importance of preserving the student’s face. They would then predict that she would select the FTA that she believes will allow her to achieve her goal with as little damage to the student’s face as possible, since she will be wary of retaliation by the student or by others. This, for Brown and Levinson, is the essence of politeness.

II.1.b Strategies and upshots

B&L list five strategies for performing FTAs in ascending order of directness. We can think of these strategies as the broad categories of politeness. The least offensive approach is to avoid performing the FTA altogether. In my previous example, the professor dealing with the disruptive student might decide that intervening isn’t worth the trouble. She will ignore the behavior, hoping that other students will also pretend that the behavior is not taking place so that she does not incur further damage to her face, which has already been damaged by the disruptive student’s actions.

A second, slightly riskier response would be to go "off the record," to say something that the student could understand as a request for a change in behavior without putting the professor on record as having challenged the student. She could tell the class something like “Isn’t the furnace making a lot of noise today? It’s hard for me to lecture well when I keep hearing noises.” If the student understands the request, he can comply under the pretense that he was never actually challenged and

411 The rest of this subsection summarizes Brown and Levinson, Politeness, 68-74.
412 Or hopefully something cleverer than this.
that he independently wanted to stop making noise. That said, he could also fail to fulfill the implicit request because he might feign to misunderstand what is being asked of him. In other cases, he sincerely might fail to notice the professor’s request. Here if he does not comply, the professor also avoids damage to her face, since she can feign that she never even made a request.

If avoidance or indirect approaches seem unlikely to work, the professor can up the ante by performing an FTA while trying the mitigate its potential damage. B&L’s third strategy is positive politeness, where an on-the-record FTA is made while indicating that the person doing the act (1) thinks highly of the interlocutor's wants and personality traits, and/or (2) shares the interlocutor's wants and/or (3) indicates that she and the interlocutor are part of the same in-group. For example, the professor could say “Hey Greg, I love your enthusiasm; can you settle down a bit so that I make sure I don’t forget anything that might be on the midterm? I want to make sure you’re ready to crush it.”

The fourth strategy is called negative politeness and emphasizes respect for the interlocutor’s freedom of self-determination. Rather than imply that two discussants are part of an in-group, negative politeness affirms the separateness and autonomy of the discussants. This can be done directly, as in “Hey Greg, I know you have a lot to tell your friend, but please hold off on talking for a bit – I just need to get through the next ten minutes and then I’ll be out of your hair”, or indirectly, via what B&L call conventionalized indirectness, i.e.: speech which has the form of a question or request

---

413 Brown & Levinson, Politeness, 70.
but is understood to be a command. To use Brown and Levinson’s example, “Could you pass the salt” would be understood as a request by all competent speakers and hearers, but one that at least nods to the freedom of action of the participant.\footnote{Brown and Levinson, \textit{Politeness}, 70. Note that this would count as an instance of the second strategy (going off-record) if there was ambiguity surrounding the statement. Contrast "could you pass the salt?" with "It's a bit chilly in here, isn't it?", which currently stands as an off the record statement but could in principle become a conventionalized indirect way of saying "Close the window." Were it to be interpreted as such, saying it would make it a greater FTA.} In my example, saying “Greg, would you mind keeping your voice down?” is clearly a request (and an order), but the interrogative form helps lessen its sting.

This is not so with the final strategy, a bald on-the-record FTA, which offers no room for ambiguity and makes no attempt to mitigate the damage done to the other’s face. Saying “Greg, stop talking” is the clearest of all requests, although it leaves a wake of symbolic destruction as Greg and those in his vicinity must struggle to perform the appropriate facework to bring about a certain equilibrium.\footnote{Notice that “Greg, please stop talking” is not as stark because the “please” is an implicit nod to Greg’s ability to choose whether or not to comply, potentially making it a case of negative politeness. Although everyone knows that the statement is a demand, Greg and his friends can act as though he is choosing to comply with a request.} Now even those near Greg can't pretend that nothing is going on and have been sucked into the interaction. We have all experienced being an innocent bystander to this type of offense. For instance, when we roll our eyes and make a "Can you believe that guy?!" face in order to express sympathy with someone who was just rudely jostled in a subway, we show that we implicitly understand the work that must be done when a perfect stranger is the recipient of a level five FTA.\footnote{Face concerns also explain in part why we hesitate to speak up for others. Most people's immediate concern, when faced with a case of subway rudeness, is to reassure the victim. Even putting aside issues of cowardice, we understand that confronting the boor would probably exacerbate the victim's new image as weak and helpless. This is why if, in this case, subway riders \textit{could} perceive the jostling as incidental, they are likely to act as if it is.} Most of us want others to see us
positively and affirm the positive image we are trying to project. Telling someone to "shut up" is the nuclear option: it clears the path but poisons the environment.

II.1.c Implications of Brown and Levinson

There are three points worth mentioning when trying to apply this to the use of deliberation as a legitimating strategy. First, Brown and Levinson’s model posits a particular type of Model Person. A B&L world is populated by rational actors who select politeness strategies by weighing face concerns against the worth of the other outcomes being sought.

What’s more, individuals in B&L's world enter social interactions not only with private interests but also with a "face" which is largely fixed before the interaction takes place. In other words, an MP has a relatively stable face which she then tries to promote and protect in her interactions. If we used this model in a theory of deliberation, one consequence would be that would behoove earnest deliberators to try hard to get to know the facets of the stable face that participants bring into the deliberative proceedings to avoid FTAs and allow the discussion to proceed smoothly.

Second, Brown and Levinson's model recognizes a tension between face-preservation and clarity. "Is anyone a bit chilly in here?" is more polite than "Close the window," but it also increases the chance that hearer will mistake the statement for an actual question and miss the implicit request to eliminate the draft. Since deliberative democrats usually seem keen on the deliberative process being a search for "truth," deliberation under the B&L model would have a bias towards the last three, and most
likely the third and fourth, options, i.e., on-the-record requests with some mitigation for the FTA.

Third, it is possible that the very fact that people minimize the impact of their statements by selecting the least threatening FTA (weighed against the importance of the action itself) could in principle discharge our duties of respect to others. One way to think about duties of respect, at least in terms of Brown and Levinson's notion of politeness, is that respect demands that we select the least face-damaging option unless our duties demand that we perform some action which requires a harsher FTA. But the fact that we take the time to weigh the options and select the least-worst FTA, and our interlocutor's knowledge that we have done so, could, in fact, communicate that we have given the interlocutor an appropriate amount of respect. I may be offended, but the sting would be lessened if I knew the offender worked hard to take my feelings into consideration. Intent counts.

I.2 Watts and the Critique of B&L

Given the long shadow that Brown and Levinson cast over the politeness literature, it would be understandable for someone trying to shore up the legitimating function of deliberative democracy to use their theory to generate new constraints on the deliberative process. As described above, one of the chief virtues of B&L is that they provide a concept of face that allows for cultural particularities and offers a

417 Since I am ultimately rejecting the use of B&L, I will not explore this in much detail other than to say that this line of thinking, while intriguing, is still problematic if the interlocutor believes that her worth or status is such that it is wrong for me to relegate her to a harsh FTA in order to get some other less important (from her perspective) goal accomplished.
sensible way of generating culture-specific dictates. It furthermore provides outside observers with a framework they can use to determine whether or not polite behavior is taking place. Finally, B&L’s concept of face – that of a twofold image of self that is relatively stable and conceptually prior to a given social interaction – seems to be compatible with the largely atomized philosophical anarchist self-image.

In spite of these advantages, the B&L model can be subject to a number of challenges, and I will discuss two of these. First, the B&L model implies that from the perspective of the recipient of the FTA, more politeness is always better than less politeness. Yet this clearly does not square with our lived experience. To see why, imagine a professor e-mailing his doctoral student in order to request the latest chapter of a dissertation. Consider the following three options:

1. “Send me the chapter.”
2. “Could you send me the latest chapter?”
3. “If it’s not too much trouble, would you mind sending me your latest chapter? If you don’t mind, of course.”

I imagine that absent additional information, most people would say that the second option is the most appropriate. The third option, while in principle more polite in B&L terms, is likely to be more alarming. “Why is my adviser acting so weird? Is he being sarcastic? I’m doomed.” Politeness is not always viewed favorably.

418 For a more extensive list of critiques, see Watts, Politeness, 98-107 and 110-116.
419 Even this is not as cut-and-dried as it might seem. Option 2 seems best all else equal. As I will argue later on, (1) contextual information might cause us to change our minds, (2) this information is frequently hard to come by, and (3) this information can emerge from the ongoing interaction
Second, it’s hard to know in B&L’s model whether our assessment as outside observers corresponds to the lived experience of the participants. Because B&L try to provide a universal framework, they are at times insensitive to the varied use of the same words, or to how the interpretation of the same word might vary over the course of an interaction. "Hey, idiot" can obviously be an insult (impolite), but it can almost just as easily be a term of endearment (polite). Yet there are many more interpretations which are hard to classify. For example, how do we classify the drill instructor who calls a soldier an "idiot", knowing that doing so (1) will make the soldier feel as though the instructor thinks she is, in fact, an idiot, and (2) will solidify the bond the soldier has with her mates, who are also subjected to insults? If politeness is tied to face-preservation, then we can either say that the instructor is being impolite or polite (since the insult may enhance the soldier’s positive face by implying that she is a full member of the group). Is the drill instructor trying to create a bond or bullying an already socially weak recruit? We as academics cannot plausibly debate how polite this complicated utterance is without recourse to how the participants actually interpret the remarks.

With these points in mind, I propose to move away from a B&L framework and instead adopt a model similar to what Richard J. Watts outlines in *his* version of *Politeness*. Watts manages to address the problems inherent in B&L but in doing so develops a theory with reduced explanatory and predictive power. Watts’ account will ultimately be useful in presenting philosophical anarchists with an acceptable
framework of deliberation, one that leaves intact the duty to deliberate without asking them to renege on their commitment to being the final judge of how to adjudicate claims, moral or otherwise, directed at them. While he may allow philosophical anarchists more wiggle room, his work has the virtue of being more correct than B&L’s model. In the following paragraphs, I will give a summary of what I believe to be the relevant parts of Watts' critique of B&L, as well the alternative framework he proposes.

Watts’ work as a whole is motivated by his deep skepticism of the viability of models of politeness which purport to allow external observers to easily describe what counts as “polite” behavior. He instead argues that any viable theory of politeness should focus on whether or not the interlocutors understand the interaction as polite. For my purposes, this can best be explained with reference to Watts' interpretation of Goffman as advocating a more socially constructed (and reconstructed) idea of face than is present in B&L, his reliance on Bourdieu's notion of habitus, and his distinction between (im)politeness, face, and politic behavior.

Watts argues that B&L misappropriate Goffman, twisting his concept of face to suit their model. Watts believes that B&L's main error lies in interpreting face as a stable construct. Watts instead argues that "face", according to Goffman, is unstable but that it contributes, in spite of its instability, to that person’s conception of self. On Watts' view, Goffman argues that different types of social interactions contain different "lines" that we can adopt, with each line containing a range of acceptable

420 Watts, Politeness, 23.
421 Ibid., 124.
behaviors. Failure to perform one's "line" correctly, or failure to choose one of a number of appropriate "lines," will, Watts believes, generally be considered "rude". Importantly, these lines change from situation to situation. Context matters. I do not have the same range of options when teaching than I do when spending time with friends. To speak of my “face” as a stable, more or less permanent entity is problematic in a life where my demeanor, perspective, and options change radically as I move from my classroom to my colleague’s office to my home.

What’s more, face can also change over the course of an interaction. I have already said that Watts indicates that deviating from expected lines is usually understood by the participants as rude (although, as I will explain soon, Watts believes that this might be desirable in some cases), but it is also possible for the expected lines themselves to shift over the course of an interaction. There may be a finite number of expected lines one may adopt at any given moment, but this does not mean that the participants are necessarily locked into an immutable framework once they have figured out which lines to adopt.

To see this, consider one of the many examples of social interaction described by Eric Berne in *Beyond Scripts and Games*, a follow-up to his canonical *Games*

---

422 Although my discussion mostly focuses on linguistic politeness, I should point out that there are many ways of being rude. As John Kasson describes genteel culture in 19th century America: “Indeed, rudeness threatened not simply in words but in the slightest expression – a gaze held too long, an insolent gesture, an overly familiar smile, an unwelcome touch.” Kasson, *Rudeness and Civility*, 115. In other words, the lines available to us are not mere scripts; they are full-bodied roles that we must play.

423 This “Wattsonian” interpretation of Goffman would be supported by studies which show that self-categorization theories of personality, which posit that self-perception is highly variable and context-dependent, seem to perform better than self-schema theories, which suggest that we view ourselves according to stable schema. See Oronato and Turner, “Fluidity in Self-Concept.”
People Play. The standard form of this interaction, which he titled *Now I’ve Got You, You Son of a Bitch* (NIGYYSOB), could run as follows. White calls Black, a plumber, to do extensive household repairs. Black provides an estimate in advance, performs the work, and hands over the final bill. The bill runs a few dollars over the estimate. White then uses the opportunity to lay into Black.\textsuperscript{424} Black's surcharge gives White the cover to indulge in righteous indignation. Surprisingly, Black also potentially receives benefits from White’s move – he can engage in his own game of "Why Does This Always Happen to Me?", acting out a role that, while casting him as a victim, gives him the opportunity to use familiar and comfortable lines when engaging White.\textsuperscript{425} What ostensibly started as a standard business transaction, complete with lines appropriate for that situation, has turned into something very different, with each participant quickly shifting to adapt to the new set of available lines.

While an extensive analysis of Berne’s example would lead me too far astray, it is still worth examining since it will prove to be relevant when trying to judge deliberative proceedings. First, the lines available to White and Black differ once White initiates NIGYYSOB. Second, White uses the *opportunity* provided by the slight overbilling to enter into NIGYYSOB. Many, Berne included, would argue that White was hoping for this opportunity.\textsuperscript{426} White did not have to react by engaging in

\textsuperscript{424} Berne, *Beyond Scripts*, 93-95.
\textsuperscript{425} Ibid., 94.
\textsuperscript{426} In fact, Berne would argue that White and Black recognized each other as co-conspirators and have mutually tacitly agreed to play this "game." On this line of reasoning, *Black* is the first mover, as he accurately sized White up as the kind of person who would react to the overbilling in the desired way. For details on what might motivate people to act in such an apparently bizarre way, see Berne, *Games People Play*, 23-40.
NIGYYSOB – he could just as easily calmly pointed out the surcharge, maintaining the roles and lines available to both parties before the "incident." In other words, not only does a Goffman-centric notion of face imply that lines and viable choices can change within a single interaction, but also that these changes are not random. One or more people may (deliberately or not) change the nature of the interaction and in doing so change the viable options available to the participants. Interactions, then, can become struggles, but rule-bound struggles. White could not have started the interaction by yelling – he needed to be given the opportunity to do so.

Watts relies on Bourdieu’s concept of habitus to explain how we can (usually) seamlessly recognize the viable lines in a social situation. Watts understands habitus to mean (in part) "the set of dispositions acquired through socialization," dispositions "to act in certain ways, which generates cognitive and bodily practices in the individual." Interactions take place within "social fields" – social organizations of space and time – where players struggle over "capital." Capital, on Watts' interpretation, is an incorporation of resources which become part of the individual's habitus. Capital can be material, cultural, or social, each with its own "marketplaces," in which each type of capital is at stake. In essence, we enter into a field and deploy (and risk) capital with the aim of acquiring more relevant capital, which we use to improve our position within the field, or that we try to transfer to other fields.

Each field has its rules, and our socialization gives us a "feel" for the

---

427 Watts, Politeness, 149.
428 Ibid. For the purposes of this discussion, I will not discuss in detail the possible interactions between fields.
appropriate behaviors in a particular field *given* our social position. We thus come to more or less seamlessly, at least in principle, interact with others according to a set of understood norms. This type of routinized behavior is what Watts calls *politic* behavior. More precisely, "*Politic behaviour*' is that behaviour, linguistic and non-linguistic, which the participants construct as being appropriate to the ongoing social interaction. The construction may have been made prior to entering the interaction, but it is always negotiable throughout the interaction, despite the expectations that participants might bring to it."429 Putting Goffman, Watts, and Bourdieu together, we can say that our habitus allows us to, among other things, recognize the relevant and appropriate "lines" that we can take in different types of interactions. In doing so, we engage in face-work, helping each other play our lines correctly. Failure to do so is considered "rude" by those engaging in the practice. It is the violation of the rules which leads to the damaging of "face," and participants will then hope for polite behavior if it helps redress the violation.430 Ultimately, interactions are both cooperative – since we help each other maintain face – and competitive, since we are constantly trying to improve our position within social fields.

### 1.3 Comparing the frameworks

This type of framework results in a number of consequences that differ markedly from the B&L model. They can be grouped into three interrelated

---

429 Ibid., 20.
430 To be clear: the “face” corresponds to the lines people chose and is therefore not necessarily egalitarian.
categories: (1) problems establishing what counts as legitimizing behavior from those outside the interaction, (2) the primacy of the politic over the polite, and (3) the participants’ unequal opportunities to shape what counts as politic behavior.

If politic behavior depends on adherence to mutually recognized “lines” which are decipherable with reference to a habitus that is itself (!) the result of embodied socialization, then those lacking in that socialization will have trouble determining what politic behavior consists in. Broad rules can appear to help – for instance, some societies are more prone than others to insisting that people form a queue when waiting to be served – but they can often confuse matters as well.

To illustrate, consider the following question examined by Watts: how should we think about how a radio host should treat those callers who call into the program? The host’s responsibilities, which the caller and the audience implicitly recognize, include being solicitous, gently ribbing the caller in some cases, keeping the conversation moving, and helping the less experienced caller through awkward conversational moments.\textsuperscript{431} So far so good, but this does not apply to all call-in shows. Other shows are more adversarial, where the host is expected to interrupt, challenge, and contradict callers.\textsuperscript{432} Is the host in this second type of show “rude”? Not by the standards of the field in which he operates. Callers and the audience expect the host to be more aggressive than the host of the other program.

The problem for academics lies in differentiating between the two. If we imagine each type of radio program as a prototype for how deliberation might

\textsuperscript{431} Watts, \textit{Politeness}, 232-47.  
\textsuperscript{432} Ibid., 218.
proceed, how are we to determine which of the two is legitimizing? It seems as though only participants or those immersed in the proceedings can judge, especially when we move from a binary classification (“rude” vs “not rude”) to a world where there are many shades of grey. The fact that participants are also the judges will present its own problems, as I will explain soon. For now, the point is that those who lack the socialization to discern proper behavior will have a tough time assessing what is going on in the interaction.

Notice that the discussion has switched from the polite to the politic. In the B&L model, politeness appears to be an unambiguous good, at least from the perspective of the recipient of the polite behavior. As I mentioned earlier, however, politeness can be viewed negatively – as snobbery, phoniness, sarcasm, a way of communicating social distance, etc. Watts’ model provides an explanation for this. According to Watts, politeness consists of a type of “overpayment” in a given social interaction, a transaction that is noted because it over-delivers. We do not think of most behavior as polite; we think of it – if we think of it at all – as appropriate. Truly polite behavior is the type that exceeds what would normally be required, and thus deviates from the expected “line” in much the same way as rude behavior.

There are many possible reasons for this overpayment and, once again, the participants are themselves in the best position to interpret the meaning of the words, although even in these cases they can make mistakes. Prefacing an interjection with an unusual “If I may...” could be, among other things, (1) an attempt to do facework in

---

433 Ibid., 32-33.
434 Ibid., 152.
order to rectify a previous offense or rude remark, (2) a way for a rich person to condescend to a poor one, (3) a way for a someone to communicate her sense of distance, due to an affront, from her peer, or (4) a way for one peer to adopt overly polite language to tease another in order to build solidarity.

The non-declarative meaning of a polite utterance cannot be established a priori, even by the participants themselves. The superficially polite statement only makes sense given a context, the same way that the meaning of "Hey, idiot!" cannot be understood outside of its context. Whether or not it's okay must depend on whether or not the participants in the interaction say it’s okay. They must determine if the notably polite behavior is politic behavior, an insult, or an attempt to pay in excess for a previous underpayment which was experienced as rudeness. This is important because situating the responsibility for interpretation at the level of the participants should be in line with philosophical anarchists who reserve the right to have final say on the moral worth of the demands that are made of them. By adopting “politic” behavior as the standard for acceptable discussion, we allow participants themselves to determine what counts as legitimizing, all things equal.

Unfortunately, all things are not equal because this model equates politic behavior with socially acceptable behavior, not moral behavior. Discussion, including deliberation, occurs within specific fields, and the composition of these fields is neither neutral nor arbitrary. Participants in the fields leverage whatever social, cultural, and material capital they have available to them to structure the field to their

435 Ibid., 243.
advantage.\textsuperscript{436} The rules and lines of a field, in other words, are anything but egalitarian and offer different possibilities based on one's existing rank and status. In fact, egalitarianism is often considered wholly inappropriate in some fields. Given this state of affairs, we can consider two possibilities: one, that the rules of the game in a field go unchallenged, and two, that they are challenged. Both cases are problematic for those who wish to use deliberation to legitimize.

The trouble with rules that are unchallenged – indeed, not even noticed until they are broken by a rude individual – is that they have been internalized by the habitus of the participants are the result of previous social struggles. Unreflective adherence to these norms is what Bourdieu calls \textit{doxa}. More specifically: "The \textit{doxa} of a field is the 'undisputed, pre-reflexive, naive, native compliance with the fundamental presuppositions of the field."\textsuperscript{437} And, as Watts says, “those with the greatest amount of capital in the field (i.e. those with power) can also use heterodoxy to change the social order of the field deliberately.”\textsuperscript{438} For example, members of the British upper classes in the late 19\textsuperscript{th} century presented a type of etiquette that was accessible to the burgeoning middle classes only to thwart their attempts at class ascension whenever possible.\textsuperscript{439}

The mere fact that a set of rules exists – a set that establishes what counts as \textit{politic} behavior within a field, does not provide moral justification to a set of rules

\textsuperscript{436} Ibid., 150-51.
\textsuperscript{437} Bourdieu, \textit{Logic of Practice}, 68. As Watts points out, the doxa is specific to a particular field. See Watts, \textit{Politeness}, 150.
\textsuperscript{438} Watts, \textit{Politeness}, 150.
\textsuperscript{439} Ibid., 43.
either objectively, as in a case where a whole society has drunk the proverbial Kool-Aid, or to lone objectors who believe that a field’s doxa is the crystallization of a set of rules that exist to support the field’s elites. If such an objector sincerely believes that politic behavior in a given context (say, in deliberative proceedings) implicitly skews interactions in ways that run contrary to natural duties that logically precede the establishment of this set of norms, then adherence to the politic behavior would entail a moral failing on the part of the objector.

What if the objector challenges the doxa? If he can do so by masterfully using the lines available to him – i.e., by being what counts as charming or impressive within the field – that’s fine, but it is precisely those people who have not successfully internalized the rules that are both unlikely to be able to pull this off and very likely to be those who are victimized by the rules.

Another alternative is to engage in rude behavior, i.e.: impolitic behavior with no acceptable (from the perspective of his audience) attempt to mitigate the FTA. There is a chance that this could disrupt the doxa of a field and lead to a change in viable lines available to participants, yet even here this disruption is far more likely to be successful when attempted by those with a lot of capital relevant to the field. More realistically, others will consider the person rude and "out of bounds," and may demand or at least wish for remedial facework, possibly including an apology, which is the subject of the next chapter.

It is hard to see how to resolve this standoff. The objector is frustrated because

__________________________
440 Ibid., 164.
the mode of argumentation is objectionable, and the only way he can express this effectively is by offending people who have deeply internalized norms which lead them to accept those modes of communication. The offended might tell the objector to "critique all you want, but be respectful about it," and this response would be both reasonable and beside the point. Engaging in "respectful" communication would mean engaging in and reproducing modes of interaction which run contrary to the fulfillment of the objector’s natural duties to engage in mutually respectful deliberation. More deliberation does not get us out of this impasse.

I.4 Fiske

I find it helpful to think of the nuances of these models in terms of the framework that Alan Fiske proposed in his *The Structures of Social Life*. So far, my discussion has implied that the broad types of interactions can vary wildly. Fiske allows us to retain the multiplicity of communicative schema while proposing a set of models (or categories) that allow us to classify patterns of interactions. This categorization gives us an enhanced vocabulary with which to discuss points of tension within deliberative proceedings.

Fiske suggests that human beings universally use four models to structure their cognitive world, to anticipate social action, and to understand and evaluate social expectations. He argues that while these models are ubiquitous, vast cultural differences can emerge when groups assign practices or interactions to one model.

---

441 Fiske, *Structure of Social Life.*
instead of another. Furthermore, these models are “atomic” in that they provide an irreducible structure by which to analyze interactions.

These four models provide a structure which informs how participants will approach all aspects of communal life within a given field. “Communal Sharing” (CS) is characterized by a strong ingroup / outgroup border and with resources being seen as communal and shared indiscriminately by group members. “Authority Ranking” (AR) is characterized by hierarchy and the belief that some members have the right to issue commands by virtue of their rank or status. “Equality Matching” (EM) occurs when participants are expected to share a resource equally or engage in a reciprocal give-and-take of roughly equal or incommensurate (and thus incomparable, at least in terms of value) goods. Finally, “Market Pricing” (MP) is characterized by understanding the participants will try to maximize their share of goods while spending as little capital as possible.

None of these structures is inherently better or worse than the others, as we can quickly see that each is appropriate in a particular context. If the habitus constitutes (in part) a feel for the "rules of the game," Fiske's typology allows us to assess the types of rules that are in place in an encounter. As with Bourdieu, the specific rules here are not necessarily fixed; they are up for contestation. Fiske’s account shows that this contestation can take place in two ways.

To begin with an illustration, consider Richard Sennett’s description of the

442 Ibid., 25.
443 For a tabular summary of these models, see Fiske, *Structures of Social Life*, 42-49.
444 Fiske, *Structures of Social Life*, 19. Fiske is careful to note that AR presupposes that those being commanded accept the legitimacy of the arrangement.
hierarchical worker-supervisor relations in Boston factories in the 1970s. Passionate foremen who occasionally blew up at the workers were respected, not reviled as one might expect. Their outbursts signaled, in part, competence and concern for a job well done. “Such outbursts became a monthly, occasionally a weekly, ritual which ended well for both sides. It might seem odd to think this rough, regular ritual as expressing civility, but it did so, as a mutual acknowledgment. 'Yeah, he blows off steam,' a machinist remarked of his foreman, 'but he's not really bad, you know?"445 These blowups were just some of the many complex interactions within the workplace ecosystem, and using an unsupplemented Wattsonian framework would make it difficult for us to make interesting statements about the interaction as a whole.

Fiske allows us to accept Watts’ admonition against arrogantly inferring what constitutes “politic” behavior in other micro-transactions between workers and foremen while providing us with a structure that allows us to analyze the types of interactions that would be acceptable to the participants. It would enable us to assess what kinds of contestation are taking place – in this case, contestations that do not challenge the AR structure that informs the rules of interaction between workers and foremen. Anyone who has had a boss has probably spent considerable time and energy testing, challenging and trying to amend formal and informal rules. These types of contestations typically do not question the AR relationships that we come to expect within many workplaces. Contestations that challenge the legitimacy of the type of framework in effect, on the other hand, are much rarer and far-reaching. Here this

could involve changes that sought, for instance, to impose worker control over the methods or means of production.

Fiske’s work allows us to consider two challenges to our assessment of deliberative frameworks. The first and conceptually simplest problem is how to choose among the frameworks when deciding which framework is most appropriate for deliberation that legitimizes.\textsuperscript{446} AR and MP are probably out from the get-go, but it is not immediately clear if a CS framework (“We’re all in this together, everyone should feel free to take the time they need”) or an EM framework (“We should make sure that every person has an opportunity to speak; no one person should have undue influence apart from the strength of their argument”) is appropriate. The choice that is made is not arbitrary; it represents a commitment to a type of worldview and a filter with which to judge future contestations and norms of behavior.

The choice is further complicated by whether or not deliberation is supposed to be contrived or diffuse.\textsuperscript{447} When deliberation is contrived, it is difficult to isolate ourselves in a bubble and ignore the markers of differentiation that we import from our world outside deliberation. Our modes of dress and very manner of being are markers of status which affect others’ reactions to and expectations of us. For example, those who have an implicit bias against people who fit the stereotype of “white trash” will not plausibly be able to ignore the symbols of ostensible inferiority when in a deliberative setting.\textsuperscript{448} What’s more, we have good reason to believe that

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{446} I am using Fiske's typology as an example. The same problem would apply to virtually any similar type of classification.
\item \textsuperscript{447} For an explanation of these terms, see Chapter 6.
\item \textsuperscript{448} Comedian Louis C.K. astutely points out that so-called “white trash” is a favorite target of
\end{itemize}
\end{footnotesize}
these biases need not be conscious for them to affect our behavior. This would be especially likely when the behavior at hand is the sum of microtransactions that are hard to describe in the moment but are picked up by an audience.

On the other hand, if deliberation is diffuse, then it is unclear how deliberative democrats can insist that legitimizing deliberation must take place in an avenue resembling CS or EM. We are all involved in AR and MP relationships, and some deliberation takes place within them. Many people feel that it is proper that their religious leaders guide much of their political discussion, either indirectly through doctrine or directly in-person. Deliberative democrats who accept a diffuse model of deliberation would then have to make a choice: accept the legitimacy of deliberation taking place in illiberal settings or embark on a campaign of massive social transformation so that all deliberation takes place within acceptable frameworks.

The second challenge is that contestation of the overall framework cannot occur by way of deliberation, at least not in the way that deliberative democrats would seem to endorse. Deliberative democrats appear to have faith in our ability to find a way to discuss matters, even contentious matters, in a manner that is acceptable to all good faith participants. This may indeed be the case when the contestation takes place from within an uncontested broader framework, i.e., from within an EM framework.

---

449 The literature on implicit bias has exploded in the last twenty years. For a canonical piece, see Greenwald and Banaji, “Implicit Social Cognition.”

450 These gatherings are not as hierarchical as military commands, but it is clear that one person has the authority to direct the discussion, both concerning content and procedure. For examples of this, see the parkour-infused Christian movie, Leap, directed by Chris Tempel.
and that does not challenge the basic framework. In such cases, contestation can be purely substantive, as in the case of policy disagreements, but it can also be procedural in Watts' sense. An example of a procedural disagreement could involve whether or not a participant is acting in an impolitic manner by claiming that people are "tuning out" when she is speaking and that her contributions are not being respected.

There are two cases where this challenge would be serious but not necessarily fatal. The first is when the objector makes clear that her impolitic behavior is the result of previous impolitic behavior that was not adequately addressed. In this case, the speaker was ignored – a breach of politic behavior – and no remedial facework was done to address this breach. Politeness would have been warranted ("I'd really like to hear more from you; could you expand a bit?") but it never took place. The second case is when the objector makes clear that politic behavior is in conflict with the EM goals of the encounter. The difference here is that while in the first case the objector was working within the framework of politic behavior, this second case involves challenging the politic behavior itself, but always from within the framework of EM.

This can quickly lead to cases that are fatal to deliberation insofar as it stops deliberation from legitimizing from the perspective of the objector. First, others may persist in politic behavior that the objector sincerely feels is disrespectful. Second, the objector may claim that what is supposed to be an EM or CS framework corresponds much more to AR. This would be the case if markers of status such as race, gender, perceived sexual orientation, socioeconomic status, or religious affiliation strongly informed politic behavior, allowing some to have access to more and better options
when deliberating. This would especially be the case if the rules or norms of deliberation were inaccessible to some people due to lack of education or "eloquence." Rules of politic behavior that rely on a middle-class habitus will systematically shut out those who simply don't know how to act "right" without embarrassing themselves.

The problem in these cases is that there is no common framework from which to draw acceptable rules of etiquette. The participants understand the propositional content of each others’ utterances, but no propositional content will make up for an utterance that is unacceptably structured or contains markers of disrespect. When the participants roughly share the rules of the game, they can negotiate the rules in real-time with more or less success. But when there is a fundamental disagreement as to what framework is appropriate, the negotiations themselves are likely to give offense to at least one party. That party will then, consistent with the duty to respect itself, have to decide whether or not to continue the discussion.

It is now worth recapitulating some major takeaways. Politeness and etiquette – *how* we say what we say – are not mere niceties, but are deeply tied to notions of self-respect. Universal human dignity implies that we are to be treated in a minimally acceptable way, but notions of differential respect, including differential self-respect, may demand more. Second, politeness is not always positive. It is an overpayment for a breach of *politic* behavior and can represent anything from a sincere attempt to repair a frayed relationship to a biting insult. As a general rule, we expect others to maintain a regime of politic, not polite behavior. This constrains the options of the

---

451 An extreme example of this would be to preface one’s utterances with a racial epithet directed toward the listener.
participants, but they nonetheless can select one of a number of “lines” to adopt over the course of the interaction.

Third, success at performing politic behavior is usually not so much a matter of explicit, declarative learning so much as an enculturation which gives us a “feel” for the game at hand. This behavior touches on what we say, how we say it, and how we physically maneuver ourselves with respect to others. It is to a large degree embodied and, while flexible, is hard to eradicate. The nature of this enculturation also means that it is difficult for those who are *not* intimately familiar with the interaction to know what is actually taking place. Is Joe being precise or pedantic? Does the lack of eye contact imply rejection, submission, or deep thought?

Fourth, while people have different backgrounds and communicative styles, we generally can come together and create stable norms of engagement. That's the good news. The bad news is that these norms, which will inform what counts as politic behavior, are not necessarily just. Participants face a number of problems here, not the least which being that those who ordinarily accept the rules will often have good reason to believe that they are owed "respectful," i.e. politic, behavior, while others will have good reason to believe that these norms should be challenged. Challenging these norms is risky, and in an unjust set of circumstances, it will be the weakest participants who will have to consider mounting these challenges. And again, here too outside observers will often not be in a situation to adequately gauge the legitimacy of the proceedings. If one person is dominating the floor, is she bullying or leading?

In sum, we need etiquette for deliberation to legitimize but we have trouble
determining the meaning of real-life behavior. It seems as though the participants themselves are the only ones in a position to know if the deliberation is taking place along lines that are consistent with each member's self-respect.  

II. Laughter

Before concluding this chapter, I would like to discuss the importance of laughter in communication with an eye towards illustrating that (1) politic behavior can contain hierarchical elements, and (2) we are typically not good at discerning or interpreting these elements. Laughter is an excellent candidate for discussion because it is a hard-wired, truly universal behavior, present across all known societies. What’s more, laughter tends to be ubiquitous within societies, occurring in many types of interactions. It is virtually impossible to stamp out; even when we do collectively proscribe laughter in cases such as (most) funerals, there are exceptions even to those rules. A well-placed quip or recollection can provoke a bout of laughter which, even if not technically appropriate, can be appreciated for its healing qualities. The point: we simply couldn’t exclude laughter from the deliberative process if we tried. We, therefore, must come to terms with its presence.

Consistent with what I have said so far in this chapter, laughter is attached to

---

452 I am not saying that each member is aware of whether or not the proceedings are violating any other member’s sense of self-respect. While the participants are in a better position to assess the quality of the proceedings, misunderstandings do occur.

453 Sauter et al., “Cross-Cultural Recognition.”

454 This applies to individuals and collectives alike. As Michael Billig says, “Joking can appear in the direst of social environments; in fact, it can provide a way of demonstrating that one has not succumbed to exigency.” Billig, Laughter and Ridicule,” 185.
both good and bad outcomes. On the one hand, we clearly seek it out – in fact, Robert Provine has found that one of the most sought-after traits by straight women in personal ads is the ability "to make me laugh." We attend comedy shows, tell each other jokes, and otherwise make funny remarks for the sake (but not only for the sake) of introducing levity into our lives and the lives of those around us. On the other hand, mockery and teasing are belittling acts that can drive victims to despair. These negative outcomes can be very long lasting; many people can still feel the sting of ridicule that took place years ago.

A naive approach would be to suggest that deliberative proceedings should prohibit “negative” laughter while encouraging the “positive” kind. This distinction has been made by some mental health professionals for whom, according to Michael Billig, “A positive experience is something good while ‘negative emotions’ are bad emotions and should be eliminated.” And at first glance, this seems reasonable. If laughter could be harnessed, the story goes, we could help bridge the gaps that worry the subaltern deliberative democrats. Good-natured jokes bring us together; hurtful jokes tear us apart, so let’s encourage the former and discourage the latter.

There are at least two problems with this approach. First, the benefits of laughter may be overblown. The consensus on the positive benefits of laughter, even “positive” laughter, rests on shaky empirical foundations, possibly the result of

456 For an entertaining and extreme example of this, see *Ridicule*, a film that depicts the strategic use of wit and mockery in the Versailles court. Leconte, *Ridicule*.
pressure to promote feel-good research and ignore the rest.\textsuperscript{458} Second, and more seriously, this binary classification of humor seems to rest either on the idea (1) that “real” laughter is a spontaneous outburst which reliably produces positive social effects or (2) that laughter, even as learned behavior, can be deployed in an unambiguously positive manner.

Neither of these is true. The notion that “real” laughter is always spontaneous is incorrect, and the story is more complicated than it may seem. It is true that babies do laugh – in fact, a normal baby laughs more than does a typical adult – and these laughs are spontaneous reactions to tickling and other types of physical stimulation.\textsuperscript{459} This quickly shifts, however, as children learn, oftentimes explicitly, what is and is not "funny." This might be done by providing cues that instances of mislabeling items count as funny, that a father who puts fruit on his head is "silly old daddy!", or by prefacing a joke with "Grandpa is going to tell you a joke," thus helping the child differentiate between serious talk (the preface to the joke) and the "funny" talk (the joke itself).\textsuperscript{460} We may laugh spontaneously in some sense as adults, but much of the “spontaneous” laughter is the result of our cultural training. Over time, we learn more nuances of the social performance of laughter and integrate laughter as a rhetorical tool. It is notable that it is much easier to laugh when others are present, both because we \textit{want} to but also because we are \textit{expected} to.\textsuperscript{461} This kind of laughter, as Billig notes, “should appear natural and unfeigned, if the recipients are to indicate their

\begin{thebibliography}{9}
\item \textsuperscript{458} Ibid., 21.
\item \textsuperscript{459} Ibid., 180-82.
\item \textsuperscript{460} Ibid., 183.
\item \textsuperscript{461} Provine, \textit{Laughter}, 45.
\end{thebibliography}
pleasure at the joke." All this indicates that there is more than meets the eye to "positive humor." Positive laughter at a joke means laughing at precisely the right time – not before the punchline, when the speaker is building up to the crescendo, and not after too long a delay. Positive humor is effective humor, and all parties must play their roles effectively for the humor to come off. Slight modifiers of this behavior – laughter that is a bit delayed, a bit too boisterous, a bit dry, can very quickly slip into sarcasm, irony, or condescension. Assessing laughter and humor thus becomes very tricky very quickly. Slightly delayed laughter, for instance, can be (1) the result of ineffective socialization of the laughee or (2) a laugh made by a very skillful interlocutor who deliberately uses the delay to indicate that the humor was offensive or lame.

That said, this fluidity also extends in the other direction: teasing, insults, and the laughter that comes with them have long been ways of demonstrating affection between friends and family. For a structured example of this, consider the practice of "the dozens." As Jerome Neu describes it, "The dozens is a practice, common especially among urban African-American adolescent boys, of aggressive, joking exchanges of patterned insults." The point of these exchanges is to demonstrate wit, make an audience laugh, and in doing so “win” the encounter. Yet although these insults touch on generally taboo topics, there are lines that are not to be crossed. Interestingly, it is when the insults come to close to the truth that a spirited encounter...
can degenerate into a fight. As long as those lines are not crossed, however, most participants seem to come out of the encounter happier than when they entered.

There are two points that I want to emphasize here. First, in each case, it is the participants themselves who determine the meaning of the laughter or action meant to produce laughter, although they do so from within a context that they understand. Even if broad rules of engagement remain stable, the actual implementation of the rules is contested and adjudicated by the parties themselves according to prior and emerging standards of judgment. For instance, "don't be too accurate" may be a standing rule in the dozens, but does an insult about the adversary’s “dumb family” strike too close to home if a sibling was kicked out of school months ago? There is no way to know in advance because the participants themselves likely do not know in advance – they must actually engage in the practice to determine if the method of engagement is the type that would make an audience laugh at the joke rather than laugh at the target. Again, it’s not just what is said but how it is said. What’s more, the target himself has a role in this. In liminal cases, the target can viably treat the laughter as appropriate or not, much in the same way that in the overcharging plumber example the customer can plausibly react angrily or not.

Second, laughter both reifies and undermines patterns of domination. For instance, research on patterns of compliments and flattery in the offices of American

---

464 Ibid., 60.
465 On a more personal note, a friend of mine answers her sibling's phone calls with "Hey, idiot" before engaging in a discussion punctuated by lengthy bouts of laughter. When discussing mutual frustrations involving another family member, however, my friend groans "Ugh, he's acting like an idiot."
companies shows a marked difference in how bosses and employees demonstrate approval. Whereas bosses tend to use direct compliments, their subordinates use imitation and frequent laughter to express approval.\(^\text{466}\) Laughter can indicate submission, a fact that may incidentally give us pause when we learn that studies show that American women laugh far more than men when on mixed-gender dates.\(^\text{467}\)

These patterns of domination need not be directly institutional. Indirect markers play a role in exacerbating or mitigating the social effects of laughter. Body language serves as a proxy for institutional power; the meaning laughter coming from someone who is hunched over differs from someone who is puffing out her chest or sitting expansively and taking up space.\(^\text{468}\) Our presentation carries information about our status, which in turn modifies how the laughter or attempt at provoking laughter is perceived.\(^\text{469}\) Our perceived level of physical attractiveness also counts – we are far more likely to defer to others and want them to like us if we perceive them to be physically attractive.\(^\text{470}\)

The point is not that laughter always implies domination, but that the meaning of the laughter is the subject of contestation and therefore impossible to categorize a priori and difficult to categorize even after the fact. Even when laughter seems disempowering, we still need to know how the participants themselves understand the

\(^{466}\) Stengel, *You’re Too Kind*, 228-29.


\(^{469}\) Of course, one's perceived status almost always plays a role when trying to influence others, even if the participants are not aware of this. For instance, Robert Cialdini reports experimental data showing that strangers are far more likely to follow a jaywalker's lead when the jaywalker is wearing clothes denoting high status than otherwise. See Cialdini, *Influence*, 227.

\(^{470}\) Cialdini, *Influence*, 171.
event. One possibility is that they do not think it’s disempowering. They might interpret seemingly domineering laughter as egalitarian hazing, for instance. Our trying to penalize laughter, in this case, would do more harm than good.

A second possibility is that there is a hierarchical relationship but that the participants are happy to be in an AR situation. Even here, we must pause before passing judgment on the laughter. It is not so easy to determine whether or not we should endorse laughter that cements this non-egalitarian relationship. AR relationships are part of life, and it’s worth remembering that AR does not imply tyranny. We can imagine a case where people simply realize that someone in the room is smart, empathetic, and articulate and defer to her on many matters. In such cases, the rules of laughter within hierarchies apply, and we can expect differentiated patterns of laughter from the participants.

On the other hand, acceptance of perceived authority relations could represent a disturbing tendency of those who are victimized to believe that they are getting what they deserve, blunting feelings that an injustice has occurred. The “Just World Hypothesis” suggests that those faced with consistent injustice lessen the cognitive load of victimhood by coming to believe that they have been treated fairly, thus allowing the unjust relations to remain intact.471 This tendency is exacerbated by a background that encourages individuals to believe that they have the power to be happy and successful, and are therefore responsible for their circumstances.472 A broad

471 Zelditch, “Theories of Legitimacy,” 38.
472 This type of messaging is especially prevalent in the self-help literature. In his classic *Think And Grow Rich*, Napoleon hill famously writes that “Whatever the mind can conceive and believe, it can achieve.” Hill, *Think and Grow*. In the recent wildly successful documentary, *The Secret,*
consensus on what counts as politic behavior – including what counts as funny and when we are “allowed” to be offended – is problematic if we believe that systematic injustice causes the disadvantaged to adapt by coming to accept the legitimacy of unjust relations. Unfortunately, there seems to be good reason to believe that this happens quite a bit.\(^473\) Sometimes it is only much later that we realize that we had no business laughing because the joke really was on us. Worse yet, we may believe that we have no recourse because the joke counted as politic behavior, and to complain would be to invite condemnation for not being able to take a joke.

Of course, what counts as politic behavior is not fixed, and the recent controversy surrounding Alec Baldwin’s portrayal of Donald Trump on *Saturday Night Live* illustrates this nicely. Baldwin’s portrayal of Trump as a narcissistic, crude buffoon has garnered much praise.\(^474\) Trump, however, has repeatedly responded to Baldwin’s portrayals by calling the show biased and unfunny\(^475\) while suggesting that he was being treated unfairly.\(^476\) Unlike his son, Don Jr., Donald Trump has given no reason to believe that he was playing along with the joke.\(^477\)

This controversy – the fact that it is even a controversy – can be thought of as the contestation of the politic in action. The derision heaped upon Trump for demonstrating anger at SNL amounts to a judgment that he is not playing the game

---

\(^473\) Major and Schmader, “Social Disadvantage.”
\(^474\) Lawrence, “Saturday Night Live.”
\(^475\) Kruger, “Donald Trump Attacks.”
\(^476\) Ibid.
\(^477\) Trump Jr., Twitter post.
correctly. It would surely be inappropriate for Alec Baldwin to launch into his Trump impression in many settings, but his portrayal seems in line with the SNL norm of mocking prominent political candidates and sitting presidents. Trump is not supposed to complain. He could ignore the joke, but he is really is ideally supposed to play along and take his lumps. Indeed, before the feud erupted into an actual controversy, some of Trump’s critics complained that Trump was benefiting unfairly from the frequent SNL coverage. By taking the skits seriously, Trump is provoking a small crisis. His “thin skin” makes the joke even funnier while simultaneously raising the stakes and challenging the norms that allow him to be mocked.

Trump seems to be the polar opposite of a man subject to the Just World Hypothesis, and this illustrates the inherent unfairness of contestation and the politic. Trump has the resources to rail against what he perceives to be unjust treatment (which may or may not include equal treatment that is insufficiently fawning) and is apparently not subject to the psychological pressures which would cause him to accept the social world – in this case, the politic world – as it is.\(^{478}\) Trump has, of course, put himself in the public eye, and thus lives an atypical life. The fact that most of us are more or less anonymous saves us from having to deal directly with some humiliating attacks, but it also makes us far less equipped to challenge the norms of politic behavior, norms which may inform our daily interactions. Power matters, and without a platform to effectively challenge the set of values that cause others to treat us

\(^{478}\) This is not the first time that Donald Trump has fought back against satire. He brought suit against comedian Bill Maher when Maher, in response to Trump’s offer of five million dollars to charity if then-President Obama released his college records, jokingly offered to pay Trump five million dollars if he could prove that Trump’s father was not an orangutan. See Ax, *Trump Withdraws.*
disrespectfully (for instance, by “playfully teasing” us), there is enormous pressure to come to believe that this is “just the way it is,” or to grin and bear it.

This may once again bring to mind Jeffrey Green’s contention (see Chapter 3) that we should think of politics in terms of masses and elites who play by different rules and so should be subject to different constraints. I do not object to this, and I think Green’s work adds nuance to any examination of how celebrities can try to influence the norms of political behavior. Yet it would be a mistake to take his characterization of “the mass” of people too far. As Green points out, most of our interactions are not with elites, and a theory of deliberation that counts on normal people talking to each other must be sensitive to how political behavior operates within Green’s undifferentiated mass.

**III. Conclusions**

All this causes problems for a theory of deliberative democracy. We will not ban laughter anytime soon, but these types of on-the-ground issues can sour deliberation to the point that one or more participants feel as though they no longer have a duty to deliberate with rankly offensive people.\(^\text{479}\) Furthermore, we cannot give precise guidelines as to how to act other than to speak in generalities like “Be respectful.” The problem is that a boss *does* think she’s being respectful when engaging in differentiated forms of communication with her subordinate. This is how she shows respect, and she has learned to expect to be shown respect in certain ways,

\(^\text{479}\) They may have a duty to pragmatically negotiate, but not deliberate in the deep, transformational way that interest deliberative democrats.
even if those ways are potentially offensive to her interlocutor. I have briefly discussed laughter, but laughter is one of many problematic facets to consider. Body language, compliments, flattery, and insults could each be the proper subject of discussion. In each case the underlying problem is the same: what seems like no big deal to an outside observer might be grossly offensive to one of the interlocutors.

Hopefully, the preceding discussion has shown that questions of etiquette are not morally irrelevant niceties. If we accept Watts’ framework, then politic behavior structures interactions in a non-neutral way, and this non-neutral structure can be subject to critique. Critiquing politic behavior in a politic manner can be effective, but it can also entail reproducing the very structures that are morally objectionable. Polite behavior is no panacea; it can at times be crucial to demonstrating respect, as when we inadvertently offend someone and wish to communicate that he is in fact valued or when we use it to make him feel more comfortable, but it can also be used to humiliate and exclude.

Discharging natural duties requires deliberation, and deliberation does not occur in a vacuum. How we deliberate matters, and there is, unfortunately, no good way of specifying what counts as legitimizing deliberation. I argue that deliberation is legitimizing when participants feel as though the deliberative process has taken place in a context where they feel respected, both in terms of universal respect as well as in terms of the differential respect they believe they are owed. In sum: politeness and

---

480 That said, I admit that they can be morally irrelevant at times, although moral considerations come up when the etiquette reinforces social structures which ought to be challenged or when the break of etiquette becomes noticed and a potential FTA.
politic behavior matter, not because they allow us to derive a priori rules of correct
deliberation, but because they are frameworks and possible tools of contestation that
allow the participants to determine whether or not the deliberation was effective.

This is not to say that this framework precludes deliberation from succeeding.
Those believe that deliberation can help legitimize the state have three reasons to
cheer. First, this type of framework is flexible and is compatible with the entire
spectrum running from diffuse to contrived deliberative settings. Proponents of
deliberation can keep whatever model they prefer. Naturally, diffuse deliberation and
Fishkin’s “Deliberation Day” will look very different, but this just means that each
method would constitute its own field with different viable lines for participants. It is
possible that one field becomes objectionable as a whole and that another method of
deliberation would be required, but most forms of deliberation at least seem
unobjectionable in principle.

Second, some politic behavior surely is unobjectionable. Many behaviors, such
as standing in line at the post office, really do appear neutral. Better still, these
apparently fair practices may well be the result of prior successful challenges by those
who were being victimized. Everyone, from a wealthy celebrity to a lowly graduate
student, is supposed to stand in line in order to speak with an attendant. Those who are
used to barking out orders in one setting may find their lines constrained in another.481

Finally, we cannot ignore the possibility that rude people can be successful in

481 I do not want to paint an overly rosy portrait, however. As I have indicated in this chapter, much of
the deference received by the powerful and wealthy is hard to notice and feels natural to all
involved. Most people are not aware at how much harder they laugh when an authority figure
makes a joke than when the joke is made by a subordinate.
changing the rules of the interaction. I suspect that this is what some deliberative democrats have in mind when they discuss the “coming together” that occurs when deliberation succeeds in spite of an acrimonious start. Not only do policies converge, but ways of communicating may converge too. Those with an optimistic bent might believe that those who are willing to exercise some courage can have their needs heard and met if they are addressing basically decent people who are committed to civility and the desire to make deliberation work.

What does this mean for deliberation? Deliberation must be done respectfully, and this respect needs to be played out in concrete terms, not just by what is said but by how it is said. What counts as “respectful” is negotiated in real time by participants who draw on established norms of politic behavior. When a breach of etiquette occurs, the participants try to smooth things over in one way or another. Unfortunately, we as observers have difficulty determining what is going and are likely to be insensitive to sincere grievances. Worse yet, even the parties themselves can have trouble figuring out the grievances if they are using methods of communication that the other finds inherently offensive, since trying to explain the offense necessitates deploying the same disrespectful means of communication.

This is why, if we assume that participants are as earnest as philosophical anarchists claim to be, we cannot impose an expansive standard of respect that overrules a participant’s judgment, thus declaring the deliberation a success and therefore binding.482 The participants themselves must determine if they have been

---

482 A thin standard of respect may be imposed. This standard would be comprised of rejecting participants’ claims if they clearly violate other natural duties. A deliberator’s complaints of being
sufficiently respected throughout the deliberative process so as to incur a duty of obligation. But, as I have stated in Chapter 6, this does not imply that a participant who believes she was disrespected cannot be justifiably coerced into law abidance. The majority may have good reason to force minorities to obey. We should accept such an event with eyes wide open, knowing that the state is *not* legitimate with respect to the dissenters.

_____________________________

disrespected will ring hollow if she is nakedly violating the rights of others.
Chapter 8: Apology

In the previous chapter, I argued for the importance of etiquette and politeness in deliberation. We do not argue in a vacuum – if the willingness to at least attempt deliberation is an important demonstration of respect, then both the form and the content must adhere to certain norms that all parties accept. I argued that some speech that appears to be substantively irreproachable may fail to legitimize because it is said in a way that does not demonstrate sufficient respect. We do not have a duty to engage in deliberation with someone who insists on ending her well-reasoned statements with "... wouldn't you agree, fatso?" or who, more realistically, intimates that she does not respect us through more subtle means.

That said, etiquette and politeness are not enough. Political discussions are contentious and are likely to offend. This is especially the case when, as with deliberative democracy, participants are encouraged to state profound convictions and allow them to be critiqued. Feelings will get hurt, and this will likely to at least occasionally lead a party to feel that they were treated with disrespect. In these cases, an apology may be in order.

Apologies are difficult to get right, but we cannot neglect their importance if we are committed to having participants believe that they can participate in deliberations while maintaining their self-respect. In this chapter I try to tease out how apologies work, their importance in other deliberation-heavy contexts, and their importance in deliberative democracy proceedings in particular. As with etiquette and politeness, apologies appear to be both necessary for legitimation while serving as
tools of potential domination.

I. The Structure of Apology

I.1. Background

I.1.a. Introduction

In this section, I present an account of Nicholas Tavuchis's *Mea Culpa*, one of the most important works on the topic of apology to emerge after the linguistic turn's heyday in the mid-20th century. Tavuchis describes what I take to be the central core of apologies: a dyadic process that takes seriously the profound difficulty we experience when we try to reconcile the rift caused by an insult. I will argue that while his account is insufficient, it is nonetheless helpful for at least three reasons. First, the main type of apology that interests Tavuchis's – a genuine, powerful, heartfelt apology between two people – is both familiar to us and relatively unambiguous. We can thus use it as a touchstone when questions become more complicated later on.

Second, this type of apology speaks to powerful, universal needs. We have all wanted to receive this type of apology – and we have all had to give it – at some point or another. Theories of deliberation that purport to take seriously the subjective experience of the participants must take these types of apologies into account. Sincere deliberation will eventually lead to real or perceived insults, and a sensitive deliberative model must accept that these insults can derail the process, sometimes in ways that are hard to detect. Heartfelt apologies are sometimes necessary to get

483 Tavuchis, *Mea Culpa.*
legitimizing deliberation back on track.

Third, most theories of apology, while necessary additions to Tavuchis's ultimately incomplete work, conceptually depend on the type of apology that he describes. All apologies follow a standard format, and that format implicitly refers to the dyadic process that is at the core of Tavuchis's investigation. The apologies that I will consider later in this chapter are parasitic on this process, even as they try to achieve different goals.

I.1.b. What apology is not

Before beginning, however, I want to briefly discuss and dismiss two types of "false" apology, as this will both prevent confusion as well as begin to tease out important features of real apologies. The first consists of the routine utterances which, while having the form of apologies, are so far removed from expressions of remorse that they are to apologies what expressions like "Please help yourself" are to imperatives: ritualized utterances whose meaning differs from what we may expect when considering only verb tense or the specific words being used.\textsuperscript{484} For example, when I tell my student "I'm sorry, but I just can't raise your grade," we may both know that I am not actually sorry. My saying so is, to use the language of the previous chapter, politic behavior that is so routinized that to offer anything less would be considered rude, given my institutional position and the basic expectations that we share regarding the behavior of people in such positions. Many students would

\textsuperscript{484} Brown and Levinson, \textit{Politeness}, 132 and 282.
probably not even notice the apology, although they would notice its absence.

The exchange with my student, like many daily exchanges, does not count as an apology in my sense because neither of us thought that the student was owed an apology in the same way that someone is owed an apology following an unexpected offense. I did not think I did anything wrong. The student (in this hypothetical case) also does not think that I did anything morally wrong – she just wants a better grade and is (or might be) disappointed with her performance. Nor do either of us think that the poor result was due to the kind of bad luck that deserves commiseration.

This type of false apology is similar to the casual apologies we utter for small offenses or for the sake of smoothing over social "bumpiness" as we walk into a crowded subway car. Whether or not the "I'm sorry"s mumbled in that context count as real apologies (as far as this dissertation is concerned) depends on the circumstance. Usually they do not. Again, these are expected responses – like my “apology” to the student – that have been accepted as mere politic behavior. No one rides crowded subways with the expectation that they have a right to be free from all jostling. We expect to be jostled, and expect others to abide by the tacit agreement that they say “I’m sorry” (or “Excuse me”) when they jostle us in a certain way. “I’m sorry” here has nothing to do with feeling sorry – I do not remember ever thinking that anyone who jostled me actually felt sorry – and everything to do with engaging in a minor face-saving ritual. On the other hand, the apologies made by someone barreling into the car too roughly and causing an unusually large disturbance do count as real apologies. The former are instances of basic (and in the case of subways, usually
admirably egalitarian) social deference. The latter are apologies because people feel unusually wronged or put out by the rude behavior.485

The second type of "false" apology that I will ignore in this chapter consists of utterances that have the form of apology but are really above-the-board insults. Take the following example. In an early episode of The Simpsons, "Insane Clown Poppy," Homer offends members of the mob who then proceed to chase him down the street. Homer, panicking, turns and yells "I'm sorry! I'm sorry!" This placates his pursuers, who stop running and agree amongst themselves that the apology changes everything, that all is forgiven, etc. Homer ruins the resolution by continuing "... sorry you're all such jerks! Ha Ha!", prompting the chase to resume.486

Homer's "sorry" is not an apology; it is a component of an insult, and this insult differs from the feeling of being insulted that occurs when attempts at genuine apologies go wrong. Homer is not incompetent, nor is he trying to avoid responsibility for his actions, nor still trying to trap the mob members into the type of paralysis we feel when a witty rival makes an "innocent" comment at our expense while preserving the veneer of civility. Homer's second apology amounts to saying "I never apologized at all!" He would consider the insult a failure if this message were misunderstood.487

485 For now, I am introducing hard and fast distinctions, but these borders will come under heavy pressure when I discuss Austin's idea of performatives later in this chapter.
486 Groening, Twelfth Season.
487 We could imagine Homer's reaction if the mobsters failed to understand his second comment. He might yell "Hey, idiots, I'm talking to you!" and drive the point home until there remains no semblance of an apology. Note that the presence of a crowd complicated matters further. When my witty rival makes a cutting remark at my expense, what really matters is that the audience knows that I was insulted. What's more, I may be dimly aware of what took place or what to do about it. I know that something bad happened, but unlike a case of a naked insult, I don't know if an angry retort will make me look even more foolish.
The bulk of Tavuchis's work describes what I call "genuine apologies." These are apologies at their finest and, as Tavuchis points out, at their most paradoxical, since they involve giving absolution as a response to a series of statements that specifically demonstrate that the offender does not deserve to be forgiven. There are three parts to an apology in this model: a call for an apology, the apology itself, and the receipt of the apology. I will describe each in turn. Taken together, they form what typically comes to mind when we think of apologies.

I.2. Tavuchis's Tripartite Apologies

I.2.a. The call

Most apologies begin with a call for an apology, an overt signal that the putative offender violated an important moral norm. As I described in previous chapters, a stable community requires its members to abide by these norms, and many offensive acts are "managed" without explicit reference to the violation. To the extent that we collectively succeed in performing facework, apologies become unnecessary. As we shall see, apologies are often awkward and painful, which partially explains why even earnest people try to avoid them. Indeed, pointing out that someone violated a norm can itself be a violation of a more important norm – the norm of avoiding the escalation of social tension – since doing so provokes a crisis. Yet sometimes group members cannot or refuse to handle a violation of social rules via implicit facework, and decide instead to risk a crisis by explicitly naming either the

488 Tavuchis, Mea Culpa, 21.
489 Ibid., 30.
violation – "You called me an idiot in from of everyone" – or their reaction to the violation – "I felt awful after that meeting; what you said really hurt."\textsuperscript{490}

Once the offender's violations have been pointed out, her status in the moral community comes into question. As Tavuchis puts it,

As a particular form of revelatory discourse, an apology is emblematic of the offender's socially liminal, ambiguous status that places him precariously midway between exclusion (actual or threatened) and rehabilitation. [...] The crucial concern of an apology is not with new rights and obligations associated with a change in social status but rather with the reclamation and revalidation of those enjoyed prior to the discreditable transgression.\textsuperscript{491}

The situation becomes rife with anxiety on all sides as now the participants can no longer work under the assumption that they share a stable and mutually satisfactory (if not desirable) relationship.\textsuperscript{492} Before the call, we assume that those we encounter are roughly on the same page as us inasmuch as we believe that they roughly abide by the same salient moral and cultural norms. When I put out a call, I implicitly put forth a number of questions. Does the offender reject the social norms broadly speaking? Does she reject this particular violated norm? Does she accept the norm in general but thinks that it does not apply to her – in other words, that she and I disagree about her social position and the rights and responsibilities stemming from it? Did \textit{I} misunderstand the norms or \textit{my} status as someone who is entitled to be treated

\textsuperscript{490} Norm violations that are not morally salient do not warrant an apology, although whether or not a given norm is salient is not always obvious. My wearing an overly elaborate hat when meeting my spouse for a private dinner may not require an apology, but wearing one when meeting her in a public restaurant might, since doing so could cause her deep embarrassment. Similarly, someone may or may not be entitled to ask for an apology for norm-violations that are morally salient but not directed toward him. A spouse who berates her wife in a public park may, after calming down, feel the need to apologize not only to her wife but to the other people in the park as well.

\textsuperscript{491} Tavuchis, \textit{Mea Culpa}, 31.

\textsuperscript{492} Ibid., 8.
differently, and if so, did I make a fool of myself?

Although the call serves to indicate that the offender's violation of the norm has provisionally placed her outside of the moral community, it is usually not stentorian. In Tavuchis's account, the call is at its core a desire for reassurance and reconciliation. If someone makes a sincere call for an apology, she is not trying to get her jollies playing judge, but rather is expressing that she was treated unfairly. It is in the service of a basic need for reassurance of her personal and social worth that she calls into question the offender's relationship with not only her, but also with her salient moral community writ large.

Once a call has been made, the offender (or her supporters) must respond. Offenders need not accept the accusation. Some calls for apology fail on their face to demonstrate any serious norm violation. "You ate your cheesecake without offering me any!" is an accusation which, special considerations aside, fails as a call, at least in most places in Southern California. In this case, if the accuser were to insist that she was owed an apology, she would be guilty of violating important norms, notably those governing false accusations and provoking needless crises.

The call for an apology is thus fraught with anxiety and ample opportunity for missteps. Unfortunately, this ambiguity and tension also permeate the apology itself,

---

493 Ibid., 7.
494 In such cases, two results are possible. Assume Jason calls on Ryan to apologize. Ryan provides an explanation that demonstrates that, for instance, Jason misheard him and that no violation took place. If Jason persists, he would either (a) be demonstrating that he and Ryan had radically different moral standards in the first place, or (b) be trying to embarrass, degrade, or humiliate Ryan. In the latter case, Ryan would now be entitled to demand an apology from Jason for deliberately using the demand for an apology as a cudgel.
as we shall see.

I.2.b. The apology itself

According to Tavuchis, an apologetic response to a call itself consists of three parts: the acknowledgment of a violation of some moral or social rule, the admission of fault, and the expression of genuine regret and remorse for the harm done to the offended party. The first two components go hand-in-hand. The person giving the apology can't simply say "I apologize," full stop. Nor can she resort to pseudo-apologies like "I'm sorry for whatever I did that upset you," "I'm sorry you took it that way," etc. She must say specifically what she's sorry for, and she must get it right. Mishandling this step risks offending or infuriating the victim further. We have all been subject to apologies that add insult to injury by implying that we are weak or being unreasonable. We have also experienced the bewildering frustration that comes with realizing that the apology has not resolved the questions that led to the call in the first place – did the offender really know how she violated important norms? And will she do so again? Is she actually that dense or is she just playing dumb?

This incidentally points to a paradoxical feature of the kind of "genuine" apologies that most interest Tavuchis. An apology fails if it contains no explanation for

495 Tavuchis, Mea Culpa, 22-23.
496 Of course, the appropriate level of specificity may vary. More general apologies such as "I'm sorry I've been such a rotten son my whole life" are more appropriate for long-standing patterns of poor behavior that encompass so many violations that neither offender nor victim can recall them all. Specific offenses, on the other hand, must be named. Notice that there is a possibility that offenders misunderstand the offended party's interpretation on this point, apologizing for a broad pattern of behavior while being oblivious to the fact that the victim has certain very specific, painful offenses in mind.
the offense. Yet the apology also fails if the person apologizing uses the explanation to excuse the behavior. It is exquisitely challenging and sometimes impossible to describe the offense without implicitly shirking responsibility. When comedian Michael Richards – best known for his portrayal of Kramer on Seinfeld – got fed up with hecklers during one of his stand-up comedy routines and launched into a racist diatribe, what kind of apology could he give? Shortly after the incident, he appeared via satellite on the Late Show with David Letterman and, after saying that his racist remarks were "crap," expressed his bafflement. "I'm not racist," he said, "that's what's so insane about this."

This attempt at apology nicely illustrates the tension between explaining and accepting responsibility. Richards's apology was widely decried as inadequate, but it is hard to think of what he could have said to improve it. Tavuchis would probably say "Not much." Yes, he could have expressed even more contrition and spent more time abasing himself. This would have signaled that he was not trying to avoid the consequences of his actions, but this still leaves the explanation of his motives untouched. We are left dissatisfied even if we take the apology to be genuine. Richards, like so many others who try to stumble through apologies, tried to simultaneously say "I did it" and "I'm not the kind of person who does this, so this shouldn't reflect severely on my character." This makes sense given Tavuchis's framework. If we believe (as does Tavuchis) that apology is at its best an attempt at re-

497 Tavuchis, Mea Culpa, 33.
498 Faber, “’Kramer’ Apologizes.”
499 And I see no particular reason to do so.
integration into a moral community, the apologizer must simultaneously reassure the audience that he understands the nature of the norm violation while reassuring them that he is unlikely to violate the norm in the future. "I was tired," "I was drunk," or "I don't know what came over me" are all variations on this theme.

Still, sometimes the explanation does justify the behavior in a reasonable way. In these cases the explanation no longer supplements the apology – it replaces the apology and constitutes an adequate response to the call. In other words, while the call for an apology may have been reasonable, the explanation demonstrates that no apology is necessary because the apparent moral breach was no breach at all.

What counts as an adequate explanation depends on local beliefs regarding free will and the extent to which people should go out of their way to gather information and avoid actions which will hurt others. "The Devil made me do it" is an attempt at an explanation. Whether or not it succeeds depends on who is saying it to whom, and in what context. When it succeeds, no apology is necessary, although it may be appropriate to express sympathy. Apologies, on the other hand, must be admissions of culpability, usually for specific offenses. To apologize is to plead guilty and ask for clemency.

The third component of the apology is an expression of genuine regret and remorse for the harm done to the offended party. There are two considerations to bear

---

500 This can apply to both parties. Amy: "I would never have done that had the Devil not made me do it! I feel awful!" Barbara: (possibly through gritted teeth) "Oh you poor thing! How awful that are in a position where you feel terrible for something that the Devil made you do!" In fact, Amy might be so upset that she feels as though Barbara really ought to buy her a cake to make her feel better.
in mind here. First, the expression must be of "genuine regret or remorse." Taken at its face, this seems obvious. Saying "I realize that I'm at fault, but I don't care" or "I realize that I'm at fault and I'd do it again if given the chance" is at best an explanation, not an apology, and at worst adds further insult. Yet, as I will discuss below, there still seem to be cases where an utterance counts as an apology even in the absence of an expression of genuine remorse.

Second, the remorse – real or simulated – must be expressed, and expressed to someone. Merely feeling sorry is not enough. Expressing remorse in one's heart of hearts, or even in one's diary, falls short. Tavuchis agrees with the precept, popularized by adherents of the linguistic turn of philosophy, that apologies are fundamentally speech acts. The utterances that constitute an acceptable apology tend to be very specific, such as "I'm sorry" or "I apologize." Other expressions, e.g., "I feel awful" or "I wish I hadn't done that" may accompany an apology but do not themselves constitute one. As Tavuchis puts it, "the offender has to be sorry and has to say so. These are essential elements of an authentic apology. Other features, for example, offers of reparation, self-castigation, shame, embarrassment, or promises to reform, may accompany an apology but they are inessential because I submit, they are implicit in the state of 'being sorry'." This is not to diminish the importance of these

502  Ibid., 22.
503  These expressions can sometimes be interchanged, but other times they cannot. For instance, "I apologize" is often more formal than "I'm sorry." "I apologize" may at times be too stuffy given the nature of the relationship and the offense; saying it may lead the receiver to suspect mockery. In other cases, "I apologize" may be helpful (although not always necessary) to signal that the apologizer understands that one offense was especially severe and calls for a temporary suspension of intimacy in order to give the apology gravitas.
peripheral parts of the apology. Self-castigation, promises to reform, or compensation may be necessary to meet the needs of the offended party. Yet, in Tavuchis's model, they are only sometimes necessary and are never sufficient when apologizing. Saying "I'm sorry," "I apologize" or some other mutually accepted expression, on the other hand, is always necessary and is at times sufficient on its own.

That said, Tavuchis' model would benefit from distinguishing further between similar expressions like "I'm sorry," "I apologize," "Please excuse me," etc. As J. L. Austin pointed out in *A Plea for Excuses*, these kinds of expressions are not equivalent, and this is so for at least two reasons. In order to make sense of this, I will briefly pause and take an Austinian perspective on apology.

An apology is what Austin calls a performative act. Contrary to constative utterances, which are true or false, performative acts (1) are not “true” or “false”, but rather “happy” or “unhappy”, and (2) “part of the doing of an action, which again would not normally be described as saying something.”505 Saying "I do" in the context of a marriage ceremony is a performative because the statement is part of the process of creating a new set of obligations; it is part of what causes someone to become married. Contrast this with "I do" as a response to "Who likes chocolate?" Here the main purpose of the statement is (usually) to provide substantive information. Here, but not in the case of marriage, someone could reply that I am lying.

Performatives cannot be false qua their performative function, but they can *misfire* or be abused.506 A statement misfires when we do not use it correctly. I cannot

505 Austin, *How to Do Things With Words*, 6-7.
506 Ibid., 15-16.
tell the priest in a marriage ceremony: “Oh, no need for me to say ‘I do’ today; I already did so yesterday.” Nor can I say “Well, I can’t be sure, but I assume that it’s me?” (although in some weddings this might be considered an acceptable joke, much the same way that sarcastic but loving couples substitute “I hate you” for “I love you”)

Were I to try these lines when getting married, I would most likely be faced with an annoyed priest who insists on asking me whether or not I do until I say the words correctly. A statement is abused if it used correctly but insincerely. To illustrate this, Austin gives as an example a marriage ceremony that is performed correctly but under duress. 507 Here I am not messing up the performance, but I am violating its spirit by doing it in spite of the fact that I would prefer not to do so.

These distinctions are worth keeping in mind when discussing apologies. Any given expression or performance carries its own rules. These rules specify the bounds within which the apology can be proffered, as well as what it takes in order to perform the apology correctly. As Austin points out, appending an excuse to an apology along the lines of “I’m sorry, I didn’t see it” works when we step on a snail, but not a baby. 508 In fact, it seems clear that the word apologize would be out of bounds in that circumstance. “I’m sorry” could help, but only if it were accompanied by additional gestures, strengtheners (“I’m so very, very sorry! I feel awful”), and/or self-recrimination (“I can’t believe I did that, how could I be so stupid?”). Note, that these additions, while necessary here, would cause an apology to fail in other circumstances. Just as how excessive politeness ruins (or corrupts) the ostensible point of etiquette,

507 Ibid., 16.
508 Austin, Plea for Excuses, 20.
going too far when apologizing can cause the apology to fail to placate the now-confused and increasingly upset recipient. The apology has failed to come off.

The abuse of apologies (in Austin’s sense) further complicate matters. As I will discuss below, public and private apologies, while clearly related in important ways, operate under different rules. Apologies are “supposed” to be uncoerced, but it seems clear that this need not always be the case. We usually want an apology by a cheating fiancee to be heartfelt, but we may not care if a corporate representative’s apology is coerced, so long as the apology is sufficiently humiliating, is accompanied by an admission of guilt, promises restitution, and so on.

Although I will not constantly invoke Austin throughout the chapter, these considerations will lurk in the background, frequently popping up to cause problems for those who consider apologizing to be a straightforward process. Tavuchis is not wrong when he says that many of us do believe that apologizing must include ritualistic words, although, as I point out below, Aaron Lazare challenges this position. Yet even if (or when) Tavuchis is right, we must be vigilant against collapsing different flavors of apology into one category. What's more, Tavuchis's basic point is sound: any particular attempt at an apology must fit a particular form. Unfortunately for theorists, this form cannot easily be specified a priori.

I.2.c. The Reaction

Although many of us spend years hoping for a specific apology, we frequently feel uncomfortable when we actually receive it. Given the paradoxical nature of
apologies described above, this is unsurprising. After all, apologies in Tavuchis's sense are difficult to give because they are *unjustified* requests for forgiveness. The recipient must decide whether or not to offer forgiveness, to undergo, as Tavuchis puts it, "a radical affective alternation" for what appears to be no good reason. Ironically, by being vulnerable and apologizing "correctly," the offender shifts the emotional burden to the offended party.510

There is thus a *double* paradox in any successful apology: the offender must offer an explanation that does not actually explain the cause of the offense, and the recipient demands the opportunity to forgive despite receiving no substantive information that should lead her to do so. In fact, the more necessary the apology, the more uncomfortable the recipient tends to be. This is especially the case when we feel betrayed by a loved one. In a satisfactory apology, the recipient's indignation is outweighed by her desire for reconciliation. The apology is finally accepted.

I.2.d. Limitations

Tavuchis's apology is thus an interaction riddled with anxiety, and this anxiety shows the importance of apology for any process involving misunderstandings and reconciliation. We want to affiliate with others and repair rifts, but we also worry that giving in too easily may demonstrate social weakness or even reflect a lack of self-worth. Yet, as I will describe in the following section, Tavuchis does not go far enough. The external pressures exercised by the audience witnessing apology are more

509 Again, if the request were justifiable, it would be an explanation, not an apology.
important than Tavuchis realizes, at least as far as my project is concerned.

Tavuchis’s account mainly emphasizes the dyadic, idealized process, and rightly so. He is trying to get to the core of apology, and he is right to say that this core must be present whenever we recognize something as an apology. Yet he presents an incomplete picture of the social pressures leading to these dyadic apologies. The recipient can appeal to social norms when issuing a call, and the alleged offender knows what types of violations would lead to “valid” calls, but these norms also bear on the recipient of the apology. The recipient knows – has learned via the same socialization that helped her understand the idea of apology in the first place – that she "should" accept certain types of apologies in particular circumstances. To not do so would be to be seen as, depending on the case, childish, overly sensitive, unfair, or timid.\textsuperscript{511} The dyadic apology is supposed to meet the needs of recipient and offender, but apologies may be "accepted" even if the recipient is no less hurt after the interaction.

Tavuchis does discuss non-dyadic apologies, but he is so enamored by the healing power of ideal apologies that he implicitly (and at times explicitly) treats deviations from this ideal as deficient.\textsuperscript{512} Tavuchis believes that apologies are fundamentally dyadic, and yet he has trouble reconciling this with his contention that the call for an apology stems from a perceived societal rule violation. Unless the breach is absolutely clear, the offended party frequently must invoke community representatives to adjudicate the matter and to co-sign the exclusion of the offender.

\textsuperscript{511} Tavuchis, \textit{Mea Culpa}, 52-53.
\textsuperscript{512} Ibid., 51-52.
from the moral community. This is what the "call" generally amounts to. Even when done privately, the offended party implicitly tells the rule-violator that if she does not apologize, the matter may be made public. When I am offended, my focus may be on my relationship with the offender, but it is mediated by my (our) community. Apologies, even personal ones, are almost always at least latently political. This will be especially obvious when apologies happen in the course of deliberation.

I.3. Apologies in Political Science

Before amending Tavuchis's framework, I want to pause and note that his individual-centric account of apology is atypical of most of the discussion of apology in the political science literature. Political scientists usually discuss apology in terms of group apologies – usually directly involving the state – and restrict their remarks to exploring group-to-group apologies, such as the US government’s apology to Japanese-Americans interned during World War 2. Tavuchis does discuss group apology in his book, including group-to-group, group-to-individual, and individual-to-group, but usually from the perspective of how it deviates from the dyadic, "genuine" apology that fascinates him so much.

While these group-to-group apologies are important, they frequently do not directly speak to the concerns of philosophical anarchists. As I have argued earlier, modern philosophical anarchists worry that accepting the legitimate state means relinquishing autonomy and/or self-respect. They are staunch (ontological)
individualists who resist group categorization. Political apologies are not *necessarily* ineffective at satisfying philosophical anarchists, but they are suspect. These apologies can be effective at restoring or affirming the dignity of a group, and thus the dignity of group members *qua* group members, but they remain irrelevant to those group members who either (1) question the legitimacy of the state based on considerations of self-respect that are not captured by the apology – in other words, for reasons that have nothing to do with their status as members of the marginalized group – or (2) reject the apology.

This rejection is itself worthy of comment. Suppose a representative of the group – often a state official in the case of apologies between nations – accepts the apology on behalf of their group. This would satisfy those group members who either accept the content of the apology or who accept that representative's authority to accept the apology on their behalf. Unfortunately, they frequently have good reasons not to do so. For example, if I am offended by my dean's comments about my academic department, I may reject an apology from the dean even if it accepted by the head of the department "on behalf of the department." I might accept that the head of the department has authority to accept the apology from an institutional perspective – my department will not be mobilized for a confrontation and future members of the department will not enter an unsettled conflict – while I simultaneously reject her

---

513 This echoes Baugh's frustrations with Wolff, explored in greater detail elsewhere in this dissertation. Baugh's intuition is correct, even if he gets the details wrong.

514 This does not mean that those who accept a representative's authority must give up feelings of rancor. It means that they must give the apologizer proper credit for having performed a successful apology.
authority to dictate whether or not I must personally toe the line and act as though the apology were successful on a personal level. The dean may be saddened but not shocked if I tell him to go to hell even after the department accepted the apology. He can expect compliance, not reconciliation.

Consider the apologies emerging from the South African Truth and Reconciliation Commission. Following the end of apartheid, many South African officials agreed to disclose the extent of their participation in upholding the apartheid regime with the understanding that full disclosure would render them immune to legal prosecution. These hearings have been widely (although not universally) lauded as being integral to the development of relative racial harmony within the country, as they ostensibly allowed its citizens to come to terms with their painful recent past.\textsuperscript{515} Some, however, have argued against the use of the TRC because they believe it to be unjust. What’s more, in many cases they ended up being irrelevant or harmful.

In short, a significant number of victims were not satisfied with the apologies. Speaking of such victims, David Crocker argues that "An act of forgiveness that ignores proper rectification results in a relationship in which at least the victim – if not the offender – feels that the new relationship is not deserved."\textsuperscript{516} Crocker goes on to argue that for apologies to establish (or reestablish) the desired relationship, they must be accompanied by punishments that both parties feel are fair.\textsuperscript{517} Crucially, it is the perspective of the participants – notably the victim – that matter here. They must apply

\textsuperscript{515} Rotberg, “Apology,” 46.  
\textsuperscript{516} Crocker, “Punishment,” 70.  
\textsuperscript{517} Ibid., 68.
their own notions of fairness when determining whether or not the apology and accompanying punishment succeed in establishing or fixing the relationship.

What’s more, the participants may believe that the event or series of events in question fall outside the scope of apology. Whether or not this is the case is itself a political question that must be negotiated. We should nevertheless not be surprised if many people conclude that it is not. As Albert Speer, Hitler’s architect and eventual Minister of Armaments and War Production wrote,

> Whether I knew or did not know, or how much or how little I knew, is totally unimportant when I consider what horrors I ought to have known about and what conclusions would have been the natural ones to draw from the little I did know. Those who ask me are fundamentally expecting me to offer justifications. But I have none. No apologies are possible.518

Let us assume for the sake of argument that most prominent political leaders representing the victims of apartheid accepted the apologies, and did so on behalf of their "constituents." Let us also assume that the majority of the victims felt as though the apologies were at least somewhat successful and that they felt as though they were treated with the respect that they believed they deserved. Surely this does not imply that the minority who felt nonplussed or, worse yet, further offended by the apologies should somehow be compelled to accept them. Tavuchis certainly could not say that they should. Apologies in Tavuchis's sense are close to paradoxes: the apologizer makes a plea, offering no justifications, and hopes (but does not have the right to expect) that the recipient will offer forgiveness. It is unsurprising that they easily fail.

---

518 Speer, Inside the Third Reich, 113.
Yet not all apologies are paradoxical. When we go beyond Tavuchis, we see that apologies can serve straightforward social needs. At a certain point, as I will discuss below, the apology will even become theater.

II. Lazare and Critique of Ideal Model

Tavuchis's account of apologies, while useful, is incomplete. He does an excellent job investigating the type of apology that most interest him – dyadic apologies between earnest agents who are trying to facilitate the offender's reintegration into a moral community in a way that respects the victim's dignity. We intuitively know that these types of apologies are powerful and speak to important, universal human needs. We also intuitively know that many apologies do not look like this. These are not the failed apologies mentioned above. They serve a purpose, but not the kind of purpose that interests Tavuchis. They are the apologies which, while not necessarily bad, somehow seem tainted. Aaron Lazare's "On Apology" gives us resources to fill out Tavuchis's blind spots, and it is to this text that I now turn.

II.1. Lazare

Like Tavuchis, Lazare believes that apologies speak to deep human needs for acceptance and respect. But while Tavuchis acknowledges that apologies frequently fall short of his ideal model, Lazare argues that apologies serve many functions and take on many forms, with no one form being somehow more genuine or inherently desirable than the others. Lazare shows that different apologies are appropriate in
different contexts and may serve a wide range of important emotional and social needs.

We apologize, Lazare says, for both internal and external reasons. Internal motives include empathy, guilt, and shame. Guilt and shame (and possibly empathy) do have a social component – for instance, we must learn (for the most part) what behaviors are shameful – but these social components have been *internalized*. Internalized social rules are, for better or worse, accepted from the perspective of the individuals that hold them, and are thus the kind of rules that philosophical anarchists would respect. Lazare says that those who apologize as a result of these rule violations “seek to restore and maintain their own dignity and self-esteem”, motives that are consistent with the philosophical anarchist project.

So far Lazare's account is in line with Tavuchis's, with these internally motivated apologies corresponding to Tavuchis's “genuine apologies”. In these models, I want to apologize because to some extent I couldn't live with myself otherwise. I don't stay up at night worrying about how my misdeed will come back to harm me; I stay up worrying about what kind of person I am as a result of having broken an important, internalized moral rule.

External motives, on the other hand, correspond roughly to what I call

---

520 An especially cantankerous philosophical anarchist might complain that internalized rules are violations of autonomy unless they are freely chosen or satisfy some objective criterion, such as Wolff's interpretation of Kant's Categorical Imperative. Having noted this possible objection, I leave it aside for the remainder of the dissertation.
Tavuchis's degenerate apologies, and are at first glance far more suspect. These are the apologies I make when I want to influence how others see me and behave toward me. If I don't apologize, I may suffer penalties other than pangs of guilt. I may thus apologize without feeling sorry about the offending behavior at all. In fact, feeling sorry about the offending behavior itself indicates that the apology is internally, not externally motivated. Unlike internally motivated apologies, externally motivated ones have the potential to be strategic and seem to undermine the logic of "real" apologies.

It is easy to see why we might view these types of apologies cynically. They include the worst examples of self-serving lies and generally smarmy behavior. This is the domain of disingenuous, disgraced politicians crying crocodile tears and trying to salvage the wreckage of their careers. We are in the realm of the business manager who, being well-versed on the latest in pop psychology literature and with an eye toward increasing productivity of his employees – or "partners", as he calls them (as

522 "Degenerate apology" is my term, not Tavuchis's.
523 Compare two cases of "letting down someone who is very important to me". The first involves a hypothetical hated boss, the second a former mentor. Both are important, but I can apologize to the former while retaining a straightforward interpretation of events that leaves my sense of self intact: "I'll apologize to the idiot because I don't want to sacrifice my career prospects. Time to start practicing my fake smile..." Apologizing to the mentor is far more theoretically problematic. Assuming that I do not feel as though I did anything intrinsically wrong, I might still apologize because (1) I want her to approve of me and to see me in a certain light, or (2) I empathize with her and do not want her to suffer. It seems that both motivations are somehow connected to my sense of identity, if for no other reason than the point of the apology is to maintain a relationship that is intrinsically important to me.
524 The apologies may be strategic at varying levels of abstraction. For instance, I may decide to adopt the maxim "Apologize whenever the person asking for an apology is socially weaker than me" because this policy usually benefits both parties. I don't feel put out by the apology, and I know that there's a good chance it will mean a lot to the recipient. I don't feel as though I have a duty to abide by this maxim, but I choose to do so anyway. Abiding by this maxim is strategic in a higher-order sense while ignoring the particularities of the situation except insofar as it helps me determine whether or not the situation should be handled by a given heuristic.
long as doing so incurs no additional financial cost) – apologizes in order to appear
more personable, to raise morale and ultimately to increase productivity and the size
of her year-end bonus.525

It is nonetheless a mistake to reject all externally motivated apologies as
counterfeit and exclusively the domain of the self-serving schemer. These apologies
can tap into human needs that are just as basic as the desire to avoid guilt and shame.
According to Lazare, these include the desire to avoid abandonment, stigmatization,
and damage to reputation.526 These needs are deeply ingrained within us. Indeed, there
is evidence that they are as part of our physical make-up as our desires for food or
shelter.527 To chastise someone for being motivated by these concerns is in effect to
blame them for being human.

Externally motivated apologies may fail to satisfy to someone like Tavuchis,
and there is, in fact, something unnerving about them. The near-universal need to
prevent damage to one's reputation, for instance, can seem quite suspect.528 How is

525 Zheng et al., “Saying Sorry.” Interestingly, these strategies may not work as well as the manager
expects.
527 For an in-depth argument in defense of this point, see Cacioppo and Patrick, *Loneliness*. The
authors claim that our need for affiliation is related to the interplay between our cerebral cortex
and our limbic system. They go on to say that once we have identified that we are not sufficiently
affiliated with a group or that our affiliation is in danger, our limbic system reacts to the stressor
just as intensely as with other powerful stressors, such as the threat of violence. The prospect of
loneliness, in other words, is deeply terrifying, and is not the type of fear that we cannot override
(for long) with cognitive gymnastics. If the authors are right, trying to think our way out of the
need for affiliation would be analogous to trying to think ourselves out of a feeling of hunger when
fasting.
528 I say near-universal because some rare people seem to be immune to this. Diogenes of Sinope
apparently sincerely lacked concern for keeping a "good" reputation, as evidenced by, among other
things, his masturbating and defecating in public. In fact, he routinely deliberately acted
unconventionally to prove a point. See Navia, *Diogenes the Cynic*, 67-68. But these exceptions
prove the rule. We are stunned when we find out that someone truly does not care about their
reputation, and we tend to admire them, our distaste for their specific actions notwithstanding.
this apology any different from that of the detested politician who apologizes for the sake of career advancement? There are at least two answers to this. First, there is a difference between human needs and human inclinations. Everyone needs food and (nearly) everyone has the desire to gorge on sugary snacks, yet we feel differently about someone who steals the proverbial loaf of bread than we do about someone who steals a bag of Twinkies "because I really like them." Our need for inclusion may fall on a spectrum, as John T. Cacioppo and William Patrick argue in Loneliness, and while this may give rise to liminal cases that blur the line between need and condemnable selfish desire, we can safely call our impulse for basic social consideration a genuine need.

The second way that "respectable" externally motivated apologies differ from the condemnable self-serving apologies of politicians lies in the theatricality of, ironically, the non-politician's apology. Externally motivated apologies are motivated by a desire to maintain one's social standing, and this social standing has important implications for how one sees oneself. When I apologize for external reasons, I use the form of apology to re-establish my adherence to an accepted social position. In fact, in many social apologies, it would be odd if my audience found out that I truly was sorry. Take for instance this account of the apology made by George S. Patton after he famously slapped two soldiers:

Following his military victory in Sicily during World War II, Patton slapped and threatened two U.S. soldiers who were patients in field hospitals and who had no visible wounds. He assumed the men were

---

529 I am appealing to what I believe to be standard moral intuition. Privately, I do not necessarily believe that the distinction is all that cut and dried.
530 Cacioppo and Patrick, Loneliness, 24-25.
suffering from psychiatric (nonmedical) conditions. Patton's physical abuse was an offense that could have led to a court martial. [...] However, Eisenhower did not want to lose his most effective field general, and so he demanded that Patton apologize to all personnel in several field hospitals as well as to the two offended soldiers. The apology was intended as a punishment to appease those who regarded Patton as an arrogant, insensitive individual. Patton, however, believed the only thing he did wrong was to displease Eisenhower. [...] His apologies were therefore insincere, but they effectively satisfied the needs of the various offended parties. 531

If one aim of Patton's grudging apology was to publicly clarify his social position and the position of the assaulted soldiers, the conniving politician, on the other hand, apologizes to hold on to or gain a position under deceptive pretenses. Patton's apology was ceremonial and depended on his saying he was sorry, not on his convincing anyone that he really felt Patton's apology was ceremonial and depended on his saying he was sorry, not on his convincing anyone that he really felt bad. 532 In this sense he is similar to a stage actor: we can have a moment of catharsis as we observe him performing his role, and he does have to perform his role with at least minimal enthusiasm, but no one believes that he truly feels sorry for the "right" reasons.

The politician, on the other hand, is using feigned – not staged – sincerity not for the reasons that motivate "genuine" apologies – those based on guilt or empathy – but rather for the sake of motivations attached to external apologies. We are usually driven to apologize by both impulses, but the conniving politician (or boss, or employee, or student...) is guilty of trying to convince us that she has internal

---

532 In fact, Patton's superiors may have been disturbed had they found out that he actually felt chastened, as this might impact his future decisions. It is possible that they would never have made him go through with the apology if they thought that he really would have felt ashamed.
motivations when these motivations do not exist. Patton can't say that his motives are external, but, although the form of the apology remains the same, he does not pretend that his motives are internal.\textsuperscript{533} The recipient and the audience know the score and are happy enough with how the ceremony played out.

Patton's apology demonstrates a crucial point: apologies can serve legitimate purposes other than those described by Tavuchis. The person offering the apology may not believe that she did anything wrong, but the potential cost of social exclusion, social penalties, or foreseeable consequences of a non-apology may be too much to bear. One need not be a Patton to offer these types of apologies. Those demanding the apology, if they are doing so because they believe that the offender has committed a real breach of important norms, may be satisfied with an apology that re-affirms these norms and the community's ability to enforce them, even if the apologizer is, in fact, unrepentant. The victim is assured the respect and social regard – however minimal – that her social role affords her, and the community is satisfied that the person offering the apology is held in check, unable to make a play for the type of social position that would allow her to commit the breach without repercussions.\textsuperscript{534} Even if the offender does not wholeheartedly endorse the norms, she demonstrates her fealty nonetheless.

Earnest or not, these externally motivated apologies humiliate the offender. As

\textsuperscript{533} Patton's apology brings to mind a cryptic remark made by Al Pacino's character in Scarface. Angrily ranting at a group of well-to-do diners at a restaurant, he says "So... what does that make you? Good? You're not good. You just know how to hide, how to lie. Me, I don't have that problem. Me, I always tell the truth. Even when I lie."

\textsuperscript{534} Indeed, even if no actual offense was given, unchecked offenses are dangerous, as they leave uncontested the offender's implicit claim that she has the social power to be allowed to violate these norms. Given enough time, this implicit claim may come to be accepted as fact.
such, they may, therefore, bind the offender to the group at the risk of violating her sense of dignity or causing her to implicitly re-evaluate her belief in the kind of social position and respect she deserves. Yet the community is entitled to signal and demand redress for rule violations. Even the offender will likely admit as much. But what happens when the offender feels the understandable need to secure her social relationships yet believes the call for an apology is unreasonable? She may accept that the demands of her community are sufficiently reasonable for her to subsume her objections (and is willing to accept some humiliation) in order to accommodate them. Yet if this is done too often, or if the calls for apology appear especially stupid or offensive, she may acquiesce while believing that this forced acquiescence has hastened her exclusion from the moral community. By re-affirming social values under duress, she distances herself from them. The only way to maintain her sense of dignity in this case is to be very conscious of the instrumental role of her apology and the alienation that she experiences vis-a-vis the community.

II.2. Critique of Tavuchis

These remarks begin to show how we should amend Tavuchis's account. Externally motivated apologies do not necessarily require genuine remorse. They mainly require proper adherence to a mutually understood ritual. These externally motivated apologies can at times be just as legitimate as internally motivated

535 It is possible that in this case, a fumbling attempt at the ritual may suffice, although, as with the case of tacit consent that I discussed earlier, failure to perform the ritual properly may signal social incompetence. We accept the apology but view the apologizer in an even lesser light than we would had they apologized adroitly.
apologies, even if they serve different (but still crucial) personal and social needs. That said, even if internally motivated apologies are not better than externally motivated apologies, they do logically precede them. Externally motivated apologies, whatever their aims, must adopt the form of internally motivated apologies. Again, this applies even if no one is fooled into believing that the professed sentiment is actually authentic.

Tavuchis says that apologies are speech acts, that is, Austin-ian performatives. Must this always be so? Lazare rightly casts doubt on this. This is perhaps surprising; generally speaking, when I'm wronged I don't want the other person to merely feel bad, nor would I feel satisfied if they demonstrated that they feel bad with sorrowful looks or averted eye contact. As Tavuchis points out, I want them to communicate an actual, bona fide verbal apology. I want them to be sorry and to say so. Sadness and offers of reparations may buttress an apology, and they may in some cases even be necessary components of one. But most of us get frustrated if we're not given the real thing, the "I'm sorry" or in some cases "I apologize." Attempts at reasonable facsimiles often add salt to the wound.

Why this bias toward only counting standard verbal formulations as real apologies? For one, saying “I'm sorry” offers a clear demarcation between the offender's expression of their emotional state and a socially recognized signal that the offended has been wronged. The distinction is essential as far as the recipient is concerned. A legitimate apology truly puts the spotlight on both parties even as it

assigns the parties different but complementary roles. A description of one's feelings, on the other hand, places the attention squarely on the offender and can turn the apology – an ostensibly debasing ritual – to a self-indulgent therapy session. Thinking of apology as speech act stops or at least mitigates the possibility of the apology becoming an expression of narcissism.

Apologies are acknowledgments of violations of certain types of shared norms. This also applies to private apologies, where community norms are salient in at least two ways. First, those engaging in private apologies may nonetheless appeal to accepted norms in order to determine when to make a call and what type of apology is appropriate under the circumstances. We see this when people use expressions like "I think you owe me an apology. I've been perfectly reasonable and that anyone else would have told you to go to hell a long time ago." Second, the parties involved in a private apology may eventually come to disagree about the content and scope of a past apology. The fact that an apology is a speech act allows both the apologizer and the recipient to ask relevant others to adjudicate and assess whether or not the apology should count or, to use Austin’s term, is felicitous.

While these concerns are legitimate, Lazare shows that apologies can sometimes take place without a speech act. These gestures, while harder to identify than a verbal "I'm sorry," nevertheless provide all the satisfactions of explicit

537 In fact, a skewed understandings of community norms can lead to apologies which we would collectively see as inappropriate. To the extent that an abusive person can keep their spouse isolated, they can convince the spouse that the abusive behavior is normal and appropriate, and that their victim's behavior deserves punishment. The victim may come to sincerely apologize for behavior that we would collectively consider unproblematic.
apologies, at least to the participants who are directly involved. The following account is an illustration of an apology that does not contain the magic words, "I'm sorry" or "I apologize":

I worked at my machine for 30 years. One day, something happened between me and the fellow next to me. Some unpleasant words were exchanged. I did not remember what was said or who was at fault, but we stopped talking. We did not speak to each other for the next six years. One day, I turned to him and said, 'I have been a damned fool,' and I stretched out my hand for him to shake. We shook hands. The grudge was over. Several workers nearby came over and asked what was going on. I said: 'I do not have to be a damned fool all of my life.' That was my apology.²³

Sometimes no words are needed at all. We can imagine many examples, usually melodramatic, where this is the case. We might think of a scene where a father angrily breaks his son's prized toy. For a long time, he resists admitting fault, fearful that doing so would erode his parental authority and sense of pride. His son – indeed, all his acquaintances – know him as a man that never admits fault. Finally, after some other event, he reflects on his relationship with his child. Years after the toy-breaking incident, he knocks on his son's door and sheepishly hands him a replica of the toy. They embrace, and all is forgiven without saying a word. We can imagine the son recounting the experience to a friend and saying "My dad apologized in his own way."

It seems wrong to say that this isn't a real apology. Something happened, and that "something" fulfilled the criteria for an apology as far as both father and son are concerned – we wouldn't bat an eye if we read in the son's diary that "Dad apologized today and it meant a lot to me." I suspect that we may hesitate to call the father's

²³ Lazare, On Apology, 142.
action an "apology" for two interrelated reasons. First, the action is not a broadly accepted way to signal an apology. Second, the meaning of the action is open to interpretation. We might accept that the action could be an apology, but it could also be something else, such as an act of pity.

As observers, we feel may uneasy because this type of apology prevents us from participating (even indirectly) in the apology. We may think that it is too idiosyncratic to be identifiable as an apology – it seeks to redress a violation of a norm, but we cannot reliably interpret the "apologetic" action as a sanctioned expression of apology. For one, even if we observe the scene, we may not even sure that the participants believe that an apology took place unless we ask them.\textsuperscript{539} What's more, even if we are assured by both participants that they feel as though a satisfactory apology occurred, we want to know that the apology was made and accepted for the right reasons. Our collective interest in policing the boundaries of acceptable behavior is stymied when there is doubt as to what exactly happened in the "apologetic" encounter.

With this in mind, we can see that the appeal of classifying apologies as speech acts lies in the seemingly universal intelligibility of these acts, at least from the perspective of the relevant community. When someone apologizes in a conventional way, any competent observer can take a stand and declare that the apology was required, or unacceptable, or over-the-top, or any number of qualifiers. This does not seem to be the case with unspoken apologies.

\textsuperscript{539} And explicitly asking if an apology took place could in some cases undermine the apology that occurred.
Yet this worry is misplaced. The father and son are negotiating an apology in their own way but are not drawing on norms that violate community standards. We need not be concerned that these apologies strip apology as a whole of its role in wider social integration as long as the parties could *in principle* explain themselves intelligibly to others. The mechanic did so, and while the explanation was useful in clarifying the event to observers, it was not necessary for the apology to succeed.

Two final remarks are in order. First, by emphasizing the subjective components and paradoxical intimacy of apologies, Tavuchis understates the importance of social position in apology. Egalitarian fantasies aside, people of different social ranks must apologize at different times and in different ways. I am not making a profound moral claim – the "must" comes from social expectations that others believe to be morally salient. Nor am I denouncing society for being non-egalitarian.⁵⁴⁰ As I have pointed out in my discussion of Fiske in a previous chapter, much of our social organization is non-egalitarian, and most of us are fine with that. The problem, at least from the perspective of deliberative democracy (which I will discuss explicitly in section IV of this chapter), is that the methods used to demonstrate respect oftentimes do so by cementing the types of inequalities that would delegitimize deliberation.

Finally, although Tavuchis notes that the person receiving the apology may feel uncomfortable, he does not go far enough. The person receiving the apology also faces pressure to "be reasonable." Although an apology is usually rendered void when the

---

⁵⁴⁰ I am not denouncing it here, anyway.
apologizer demands that it be accepted, the victim may quickly become the villain if others feel that she is holding on to a grudge for too long. This means that the offender is not the only person who risks humiliation. Recipients may be coerced to let bygones be bygones, even when doing so fails to meet their needs. Those observing the apology are not mere bystanders. They have their own interests, and these interests may or may not align with those of either party. Regardless of the motivations for the apology, both apologizer and recipient may be participating in a ritual even though they believe the process to be unjust.

In this section, I tried to add nuance to Tavuchis’s powerful account of apology. We apologize for all sorts of reasons, and the fact that these reasons may be externally motivated does not lessen their import, either for the direct participants or the community at large. Some apologies are successful when they don’t express genuine remorse, and not all apologies must be speech acts. Most crucially, apologies are not dyadic processes in which observers have marginal effects. Many apologies are performances, and these performances are as much done with the audience in mind as with any desire to achieve genuine reconciliation.

III. Restorative Justice and Apology

III.1. The Importance of Apologies in Restorative Justice

Before examining the implications of apology in deliberation and legitimation,

541 Lazare, On Apology, 233.
I want to discuss its place in restorative justice practices. These practices resemble deliberative democracy in many important respects, and insights taken from the former can help highlight problems in the latter.

For the sake of having a clear starting point, I begin with a definition offered by Joanna Shapland: “Restorative justice is a process whereby all the parties with a stake in a particular offence come together to resolve collectively how to deal with the aftermath of the offence and its implications for the future.”  

Restorative justice practices are approaches to criminal justice that, while quite varied in their particulars, all object to what they take to be the over-institutionalization of the judicial process. Proponents of restorative justice believe that standard court proceedings disempower those most affected by the alleged criminal acts, whether victim or offender. Victims – people who typically need compassion and understanding – move through an institutional process which, far from attending to their actual needs, can exacerbate their pain by subjecting them to harsh questioning and forcing them to relive the crime without adequate avenues to express their anger and hurt in a constructive way. As a result, a surprising number of victims who win cases find themselves dissatisfied with the outcome.

Restorative justice proponents hold that there is no one-size-fits-all solution to satisfy parties involved in a judicial dispute. The needs of victims differ – some simply want restitution, others retribution, still others a chance to engage with the offender.

543 This is perhaps most evident in cases of alleged sexual assault. See for instance Nowrojee, “‘Your Justice’,” 132.
What is clear, however, is that courts frequently fail to take these needs into account. In fact, some argue that the victim's process through a court case amounts to a second victimization in which feelings of powerlessness are reinforced, not diminished.\textsuperscript{544} What's more, the victims are not the only ones who lose out – offenders tend to become even more recalcitrant as a result of their engagement with the criminal justice system.\textsuperscript{545} Not only are offenders more likely to re-offend, but they are also more likely to feel justified when committing future crimes.\textsuperscript{546} Finally, the community as a whole must bear its share of these costs while typically having little or no say in the proceedings. By trying to impart impersonal justice, the courts risk alienating everyone involved in the proceedings.

Restorative justice advocates seek to use deliberation to reform or replace the criminal justice system with a process that would yield better outcomes from the perspective of the victim, offender, and community. They seek to bypass the standard rehabilitation vs. retributivism debate – should we help offenders change their ways or should we give them their just deserts? – and in doing so, minimize the role of formal institutions and abstract justice in the judicial process.

At its core, restorative justice consists of bringing relevant parties together and allowing them to discuss the offense and its impact in a relatively unstructured way. The discussion can be painful, but restorative justice advocates generally celebrate these expressions of pain. From their perspective, the free exchange of views – not

\textsuperscript{544} Zehr, \textit{Changing Lenses}, 23.  
\textsuperscript{545} Johnstone, \textit{Restorative Justice}, 88.  
\textsuperscript{546} Zehr, \textit{Changing Lenses}, 41.
merely statements of facts, but expressions of hurt, hope, desires, and so on – are not only permitted but *required* for justice to take place. This type of interaction allows participants to (1) articulate their immediate needs, (2) agree to satisfy each others’ needs – e.g., with unorthodox restitution or with a particular type of apologetic behavior, and (3) meet additional incidental needs through the very process of deliberation. For example, victims and offenders may, for instance, come to humanize each other through the interaction, relieving the victim of thoughts of an all-powerful oppressor and the offender from a sense of alienation that may have contributed to the offense.\(^{547}\)

The upshot is that the restorative justice process, if done right, legitimizes the result of the interaction because the participants are able to have a meaningful say in the outcome and were able to do so in a way that maintained their dignity, autonomy, and self-respect. Participants who have gone through the process have found the process to be fairer because, compared to formal proceedings, here they were “more empowered to control the discussion.”\(^{548}\) This is exactly what deliberative democracy tries to achieve. What's more, most restorative justice advocates at least mention in passing that the restorative process promotes social cohesion, another ideal associated with deliberative democracy. This desire for social cohesion will become especially explicit in my discussion of John Braithwaite's model below.

Restorative justice advocates believe that giving and receiving apologies provides the glue for this social cohesion. A typical ideal scenario runs as follows. The

\(^{547}\) I should re-iterate here that I am describing an ideal encounter, not a typical one.

\(^{548}\) Braithwaite, *Restorative Justice*, 78.
victim and offender agree to use restorative justice practices and meet under the supervision of a state official or community representatives.\textsuperscript{549} The victim is given a chance to express her pain and ask the offender questions.\textsuperscript{550} In some models, such as Braithwaite's, members of the community may be permitted to do the same. The offender listens and responds to the questions, expresses herself, and a more or less unstructured dialogue takes place.\textsuperscript{551} The dialogue culminates – and remember, this is an idealized version of the process – when the offender offers a sincere and heartfelt apology to the victim (and if the community if applicable), who, having understood and empathized with the offender, accepts the apology. The two form a type of bond in which each person is necessary for the other's healing and full integration back into the community. Both offender and victim feel supported and understood. As a result, they both willingly negotiate an outcome that takes their needs into account. A desire for reintegration, not an impulse toward retribution or "fixing" the offender, guides the search for a mutually acceptable outcome. If retribution is important, it is because it

\textsuperscript{549} The vast majority of restorative justice proposals and practices that I have come across situate the process alongside a more formal criminal justice system. In general, both victim and offender have the option of abandoning the restorative process and switching to the institutionalized track at any point, a precaution that seems necessary to prevent the violation of either party's rights. It is not always clear if this two-track model is supposed to be a permanent solution or a necessary artifice that is to be abandoned once the utopian vision of restorative justice is firmly entrenched.

\textsuperscript{550} It turns out that an overly rigid emphasis on proper procedure and abstract justice prevents us from noticing that in many cases, what people sometimes really want is to make sense of the situation, to know "why me?", etc.

\textsuperscript{551} The state official or other facilitator does, however, ensure that the conversation does not go completely off the rails. That said, "irrelevant" conversation may in some cases be fine or even welcome. A victim and offender who exchange general opinions about irrelevant topics may be forging a closer bond. In fact, it's not always clear when the facilitator should intervene. A very emotional victim may need to "get it out of his system," and an offender who is being emotionally pummeled may (1) deserve it, and (2) actually be having her needs met. Feeling awful may be an indication that she is facing up to the pain she caused for the first time. For more on the latter, see Morris and Young, \textit{Reforming Criminal Justice}, 262.
meets the needs of one or both of the participants.\textsuperscript{552}

The extent to which this idealized version plays itself out in practice is up for debate. Proponents of restorative justice argue that it happens more often than we might expect, so long as the process is structured appropriately.\textsuperscript{553} Many further argue that the success rate would be even higher if restorative justice practices were more fully integrated into (or replaced) the standard criminal justice system. Without a full commitment to deliberative practices, these advocates say, it should come as no surprise that some attempts at restorative justice fail.

And some attempts at restorative justice clearly \textit{have} failed. Take for example the restorative justice program instituted for young offenders in Britain as described by Gwynn Davis in \textit{Making Amends}.\textsuperscript{554} Qualified young offenders were given the option to participate in a pilot restorative justice program and bypass the criminal justice system. These young offenders, under the supervision of a state official, met with victims who chose to participate in the project, and offered an apology. This resulted in a number of awkward encounters that left no one satisfied and that did not seem to affect rates of recidivism.

These apathetic encounters seemed to have all the qualities of a bad date. They were disappointing, nervy affairs that both parties felt obligated to see through until the end. The offenders seemed unmoved – there were no powerful moments of self-
discovery or bonding. The victims seemed to fare even worse. They mostly believed that the offenders were (1) insincere, and (2) using the restorative justice process to avoid harsher sentencing. The apology, supposedly as a sincere expression of remorse, came off as calculated. As a result, many victims came away from the encounter feeling foolish for expecting some sort of positive reconciliation.

The supervising official, there to facilitate conversation, arguably made the situation worse still. Victims felt as though they did not even have the luxury of angrily denouncing the apology as insincere; many reported feeling pressured into accepting the apology. In other words, they were thrust into a role where they were required to play the role of the generous "forgiver." They thought that refusing to accept the apology would result in being seen as unreasonable or "difficult" not only by the offender but by those who set up and oversaw the process.

Victim and offender were put in a situation where they were "supposed" to connect on some meaningful level and, of course, utterly failed to do so. Yet the presence of the community representative forced them to at least pretend that the encounter was a success. People who might normally balk at returning cold soup at a restaurant were told that they could speak up and refuse to accept the apology in highly stressful and awkward circumstances. Unsurprisingly, many did not refuse but nonetheless expressed dissatisfaction with the interaction when interviewed after the fact.

That said, even these “failed” apologies had their bright sides. Some victims

---

555 Ibid., 13.
556 Ibid., 65.
wanted a chance to express their anger and hurt – indeed, this is one of the most pressing needs that many victims seem to have\textsuperscript{557} – and did so during the encounter. What's more, even if many victims were thrust into an undesired role of "victim who graciously accepts an apology," they also reported feeling some satisfaction at playing a part in the socialization of the young offenders. This is important. In interpreting the interaction along the lines of "Adult tut-tutting naughty child," the victim can re-establish herself as a valued member of her community. This is in many cases crucial for the well-being of the victims, who (1) feel diminished by their status as a victim, (2) want reassurance that they are supported by their peers, and (3) want to feel that they are not socially diminished by their loss.

These needs are not fulfilled when one feels coerced into accepting an apology but can be met when the victim is able to accept her socially sanctioned role. To use the language of transactional analysis, the insecure victim is, by virtue of her victimization, thrust into the role of "Child," with her support system being a well-meaning but potentially humiliating "Parent." By playing the role of "socializer who is helping re-integrate the young offender," the victim re-establishes herself into the collective "Parent" fold, thus abandoning her "Child" standpoint and participating, by affirming the dominant social values, into the shaming of the "Child" offender. By helping to "tut-tut" and ultimately forgive the offender, the victim may ignore her idiosyncratic needs for genuine connection and expression of pain but at least gains confidence that she is socially powerful.

\textsuperscript{557} Zehr, Changing Lenses, 22-26.  
\textsuperscript{558} For a more in-depth description of these roles, see Berne, Games People Play, chapter 1.
We are now quite far removed from the ostensible purpose of the apology – a heartfelt connection designed to heal emotional and social rifts between the victim, offender, and community. Yet if this is a failed apology, it may be more beneficial to all parties than the feasible alternatives. The "failed" apology might end up being a legitimate externally motivated apology, one that reassures the victim of her social standing, allows the offender to escape certain types of punishment (in exchange for ritual degradation), and ultimately reassures all parties that everyone knows their place.

There are two upshots here. First, even if we accept that the real benefit of these apologies lies in victim empowerment, the interaction simply wouldn't make sense (or, more precisely, would be radically altered) without the background fiction that this is a “real” apology. We must go through motions that might fool a naive observer into believing that real apology took place. Second, if the apology is of this “failed but beneficial” type, we must accept that whatever re-integration takes place will be done because the offender has been humiliated. Offenders are re-admitted as “junior members” and socially demoted, or their existing place near the bottom of the

559 To illustrate, take for instance the famous "Would you like to come upstairs and look at my etchings?" line that one potential romantic partner might ask another or, more commonly, invitations to "come upstairs for some coffee." We normally assume that these lines are used in order to give both parties plausible deniability – the recipient of the offer can refuse it without risking offending the date, and the person making the offer can deflect a negative response by telling herself, the partner, and others, that she was just offering a cup of coffee. Yet at other times it is absolutely clear that both parties want to have sex, yet the offer is still made. This is because although in such a case "Would you like to have some coffee" is substantively identical to "Let's have sex," nakedly asserting "Let's have sex" changes the meaning of the encounter. It means accepting a new kind of identification for the sexual relationship, that "we're the type of people who just straightforwardly say 'Let's have sex.'" This switch contains a host of implications of how the parties interact in non-sexual encounters. Similarly, allowing victims to vent without there being the "specter" of a real apology in the background has potentially far-reaching implications.
social hierarchy has been cemented as a result of the interaction. Before moving to a model of apology which attempts to collapse the distinction between internally and externally motivated apologies, I would like to conclude this subsection by talking about the role of time in apology. The restorative justice examples that I have given here, being concerned with questions of judicial proceedings, feature cases where the offense has not taken place very long ago. In some ways, this is appropriate for a discussion on deliberation – at the extreme, I am concerned with offenses that take place in real time during the deliberative process. Yet on another hand, the discussion fails to adequately consider longstanding disputes. Failing to consider the importance of these disputes means committing the fallacy of assuming that although deliberation participants have differing preference profiles partly as a result of their cultural background, they nevertheless enter into discussions as blank slates, bereft of grudges that cannot be resolved by changes in resource allocation. Examples of this abound, but I will discuss one in particular: the 1921 Tulsa Race Riot and its aftermath.

In 1921 a race riot took place in Tulsa, Oklahoma, a racially divided town where lynchings sometimes occurred. The bare facts are as follows. On May 31, 1921, a newspaper reported an alleged attempted rape of a white woman by an African-American named Dick Rowland. African-American residents feared that Rowland would be lynched, and a group of them, including a large number of World

560 I want to stress that it is possible that some people may feel, under some circumstances, that they should occupy a lower social rung.
561 Brophy, Tulsa Race Riot Commission, 237-239.
War I veterans who chafed at the servility expected of them upon their return from war, went to the jail to offer their assistance in protecting the prisoner. The sheriff convinced them to leave but did not disperse the angry white crowd that was gathering. The African-American residents returned that evening, whereupon a police officer attempted to disarm an African-American man. Someone shot, and chaos ensued. A white mob looted and set fire to buildings in the African-American community, killing some people in the process. The city council then passed a zoning ordinance mandating that the rebuilding process use fireproof materials. This priced out the former residents from their now destroyed neighborhood, which then fell under the control of white interests.

The legacy of this riot continues to this day. While everyone agrees on the barebones facts described above, there are competing narratives as to what happened in 1921. Many members of Tulsa’s African American community hold the view that the rioters were not part of an unruly mob, but rather part of a planned conspiracy, with roots in the government, designed to drive African American residents off their land with no compensation. 562 Their narrative also cements their identity as victims in conflict with an “outside, oppressor community.” 563 Many members of the White community, predictably, have a different narrative. Some feel that the riot was sparked by unruly veterans who stirred up trouble “defending” a man who was safely in police custody. What’s more, many simply want to ignore the riot as a sordid event that took

562 Ibid., 239.
563 Ibid., 240.
place in the distant past and wish that African-Americans would “get over” it.  

It is in this context that the Oklahoma legislature established a commission to investigate the causes and effects of the riot. The commission, which began in 1997 and reached its conclusion in 2001, predictably did not settle the matter. Much of the evidence was vague or no longer available, leading to a story which was sound from a historical perspective but ultimately unsatisfying to a broader audience. The legislature’s reaction to the report was to avoid it, other than to award medals to Tulsans who had survived the riot, as well as to release an apology of sorts, statement that acknowledged the moral culpability of Oklahomans and that acknowledged that some city officials did not perform their duties adequately.  

This case presents us with numerous lessons regarding the use of apology in deliberative contexts. First, and most obviously, time does not heal all wounds. Many Tulsans, even those who were not alive in 1921, hold grievances attached to the riot and its aftermath. People do not always “get over it”. Second, the older the case, the harder it is to agree on facts. This is dismaying, since most accounts of apology insist that the apologizer must apologize for the “right reasons” (from the perspective of the recipient) in order for the apology to be felicitous.  

Third, the alleged offenses here did not necessarily turn on matters of personal affront – after all, many who wanted a more thorough apology, complete with reparations, were not alive at the time of the riots. What counts is the interpretation, and how different interpretations are used to inform the identity of present-day

564 Ibid., 241-242.
565 Ibid., 244-245.
Tulsans. Unfortunately, the commission was not able to craft a version of history that presented a story that all sides could endorse and accept.

An important upshot is that people will disagree about whether or not an apology is in order, and in many cases it is difficult to say that one side is not entitled to an identity which, for instance, makes them a victim that deserves an apology. Some interpretations are probably healthier than others – hence the well-known saying that “resentment is when you take the poison and wait for the other person to die.” Yet we cannot insist that someone is wrong for sincerely holding an interpretation of their self and their relationship to an event, and wanting an apology. Nor can we call someone wrong for claiming victimhood by virtue of their holding a shared identity with those who were victimized in the past. This does not mean that those people are necessarily entitled to reparations – none were given in this case, rightly or wrongly – but we do identify as members of a group, and if we believe that “our” group has been disparaged, a lack of an apology makes the deliberative process very difficult, if not impossible. To say that these groups are part of the same political group in spite of their differing and antagonistic understandings of the connection they share raises questions about whether or not they have a sufficiently shared identity to be “political” in the same way that group members are political with respect to their own group.

III.2. Case Study: Apologies and Braithwaite

There is a possible way to use humiliation to collapse the distinction between

566 M.T. Sponsorship Guide to 12-Step Programs, 134.
external and internally motivated apologies. John Braithwaite's restorative justice proposal does this by doubling-down on the externally motivated causes of apology. In *Restorative Justice and Responsive Regulation*, Braithwaite goes beyond accepting that deliberative practices may be formulaic. Unlike those who merely say that the ritual of apology offers side-benefits, he proposes a ritualistic system which deliberately shames the offender, ultimately converting an externally motivated apology to an internally motivated one.\(^{567}\) Braithwaite accepts that offenders are unlikely to offer genuine apologies as a result of self-reflection or even as a result of hearing the victim's pain, and thus advocates that we organize “restorative” proceedings that shame offenders in order to break them down.

Unlike the unconvincing apologies offered by British youth in Davis's account, here the offenders really are moved to apologize for reasons other than coolly calculated self-interest. In this sense, Braithwaite's model seems true to the ideal model of apology: the motivation is internal. In another sense, however, something seems to have gone very wrong. Whereas in the original model the apology comes about as a result of a meaningful empathetic connection between victim and offender and is the result of a genuine understanding between them, here the offender's compulsion to apologize comes from a fear of social stigmatization which is experienced as a powerful form of guilt.

Social stigma is one of our most fundamental fears, and evolutionary biologists have traced a desire for at least basic social acceptance to a basic biological need.\(^{568}\)

---

Braithwaite's model takes advantage of this by turning an apology into a disguised act of groveling. To his credit, he is unabashed about this. He believes that heightening feelings of shame will ultimately lead to a heightened desire for re-integration into the moral community, tighter allegiance to this community once the offender has been accepted back into the fold, and lower recidivism. This model is therefore far removed from an investigation into the truth of each participants' interaction and is instead a celebration and re-affirmation of the dominant group's moral norms.

While I find Braithwaite's model repugnant, it is not wrong. His model tempts us to try to eat our cake and have it to, to try to get the benefits that Braithwaite lists – better rates of reintegration, stronger adherence to existing norms, emotional support for victims, ultimately less harsh punishments for offenders (relative to what they would face if they were processed through the normal criminal justice system) – while actually believing that the “apologies” offered are the same as Tavuchis's idealized apologies. If his model succeeds, it is because Braithwaite offers a perverse inversion of the transformative powers of deliberation.

As I discussed earlier, some proponents of deliberative democracy argue that deliberation helps modify preferences, allowing participants to converge on a set of beliefs. Braithwaite's participants also converge on a set of beliefs, but this is because of the power of humiliation masked as deliberation. A successful encounter here would be one in which a recalcitrant offender came to adopt a belief as a result of trying to cope with the prospect of having basic needs gone unmet. The result may be an

internally motivated apology free of "theater," but we are now far, far removed from ideal deliberative models steeped in concerns for egalitarian interactions. Like most restorative justice advocates, Braithwaite wants the process to be transformative. The question is whether or not we like how the transformation takes place. There is a world of difference between Tavuchis's spontaneous apology and a choreographed apology that deliberately brings the full weight of social pressure to bear on the offender.

Braithwaite represents the ultimate departure from Tavuchis's type of apology, but by insisting that we wield social pressure like a bludgeon, he brings us full circle and leaves us with an offender who, if the process has worked correctly, has just gone through an emotional hell and who sincerely wants to "make things right". By getting through to the offender in this way, Braithwaite seeks to force offenders to become their own tormentors and to seek absolution at the hands of the victim and community. This type of interaction, even if it results in greater satisfaction after the fact, seems like a nightmarish inversion of the respectful discourse desired by deliberative democrats.

IV. Apology in Deliberative Contexts

People value apologies. Apologies are ubiquitous, serve important subjective and social purposes, and vary in form and content according to the particulars of the situation. In all cases, they refer to a core form of dyadic apology. It is now time to see how these insights apply to deliberative democracy. The conclusions will be tentative.
I cannot avoid this – not only are apologies multifaceted but, as I discussed in an earlier chapter, so is deliberation. Different types of deliberation have different criteria that must be met for the deliberation to be legitimizing. Yet while instances of apologies and deliberation are idiosyncratic, we can make a few general statements about the interaction between the two. The common theme that runs through what follows is that apologies, while necessary to demonstrate respect, can do so by cementing the types of inequalities that delegitimize deliberation.

If we consider deliberative procedures along a spectrum ranging from the personal to the impersonal, schemes that emphasize deliberation among friends, close acquaintances, or members of a tight-knit group – often in one-on-one settings – will likely demand heartfelt apologies. In those cases, deliberation is tied to a construction of identity that requires preexisting bonds of affection, loyalty, solidarity, and/or trust that allow participants to deliberate in a way that goes beyond mere negotiation. We trust those we know and like; we don't usually trust strangers enough to initiate calls for intimate apologies. An insincere apology in tight-knit settings runs the risk of alienating members and preventing them from participating in the deliberative process in a meaningful way.\(^{570}\) In other words, ineffective apologies – including the type we might give a stranger – prevent hurt participants from adopting the trusting mindset that legitimizing deliberation requires in this context. Heartfelt apologies reestablish the kind of vulnerability that serves as the bedrock of close relationships.

Yet these sincere apologies may be ineffectual or even harmful if used in more

\(^{570}\) By "meaningful" I mean that while they can still participate substantively, they will be less likely to do so and, if they do, are unlikely to view the outcome as an expression of their input.
impersonal settings. In such cases, the publicity of the apology may be its most important component. The person who is rudely interrupted in a citizen jury is usually not looking to re-establish intimacy with the offender or the group since there was no real intimacy – just civility – to begin with. Here a public demonstration of recognition of wrongdoing is in order. A private apology proffered during a break would not work. What's more, an apology that emphasized genuine feelings of shame would likely embarrass everyone involved. Whereas in "close" deliberation the apology serves to reestablish intimacy, in formal deliberation it maintains an appropriate distance and removes the venom from an offense committed between strangers or mere acquaintances.

"Face" is important in each case, but in crucially different ways. In the first case, the apology serves to reassure the recipient that the partner giving the apology values the intimacy of the relationship and that she wishes the re-establish the pre-existing social dynamic which, as we have seen via Fiske, may or may not be egalitarian. In the second case, the apology serves to ratify the social worth of the recipient and, in doing so, at least temporarily lowers the social standing of the apologizer or cements her weak social position. At the highest levels of institutionalization – apologies between states – the value of apologies stems almost entirely from their ceremonial, symbolic, or restitutive character. This import of this type of apology is that it is, as Tavuchis puts it, "on the record speech," where clarity and precision are emphasized over the actual experience of sorrow.

---

571 I discuss the importance of face in Chapter 7.
Even with this distinction in mind, it is difficult to come up with actionable rules to guide behavior in particular cases. Apologies are messy, and it is not enough for us to say that a ceremonial apology is in order. We should be wary of trying to determine precise rules of apology, i.e. rules that observers can refer to in order to declare on party reasonable and the other unreasonable in the event of a dispute. To put the issue in the context of my overall project: if the idea is to use a duty to deliberate to convince philosophical anarchists to participate in a legitimizing procedure, and if a philosophical anarchist sincerely feels that she has been disrespected and cannot endorse the procedure without receiving a certain type of apology, then the participants do not have recourse to a recipe that would necessarily and reliably placate the (earnest) philosophical anarchist.

Apologies are tied to our sense of dignity as well as to our beliefs about what we deserve and what we can (and cannot) put up with as members of a political community. Because our subjective rules on this point are to at least to some degree idiosyncratic, outsiders cannot always determine when rule violations take place. This interplay between the subjective and the normative creates additional complications for the use of apology in deliberation. I describe two such complications below.

IV.1 Apology and Coercion

The relationship between apologies and coercion is problematic. As we have seen, apologies can be coerced, and while coercion voids purportedly internally motivated apologies, it may at times leave untouched or even augment the power of
externally motivated ones. It's nice if the offender has an actual change of heart, but the show of social solidarity in support of the victim is what really counts in these cases. That the offender implicitly feels the need to apologize for fear of social retribution goes a long way toward showing the victim that her place in the social order is secure. Explicit coercion powerfully highlights the point by re-affirming the victim's relationship with her allies and community at large. Yet, as mentioned above, the victim can also be coerced into receiving an apology even when she finds it inadequate or even infuriating. We can coerce receipt inasmuch as we can force someone to listen to an apology and give an indication that the apology has been accepted regardless of how they feel.

Many siblings have experienced as children a situation in which a parent coerced both parties to apologize, and to do so in a certain way. Parents socialize their children by forcing them to act as if they were performing a genuine apology. Jane may be forced to apologize to Sue, who may then be ordered to "forgive" her sister. The fact that neither sister wanted to take part in the process does not make the interaction meaningless – perhaps the rest of the family can now breathe a sigh of relief and agree that it's time to "move on." What's more, both Jane and Sue are now learning a valuable tool they can use to diffuse hostile situations and develop intimacy in the future. Yet for all the good that this coercion does, the meaning of such an apology can vary considerably for both Jane and Sue. In some cases, the coercion is not merely annoying – it is humiliating and alienating.

We need to ask whether or not the apology is taking place between members
with secure positions in the group. If Jane (the "offender") is an outsider – as was Cinderella under the tutelage of her wicked stepmother – the coerced apology will serve at least one of two purposes. First, it may be used to humiliate her further and mark her as an outsider that can be pushed around. Second, it could serve to include her in the group, usually in a subservient position. In other words, the apology, in this case, would highlight either her liminal status or inferior position with respect to the group. The basic (but not the only) idea here is that consistent humiliating interactions will lead both parties to implicitly or explicitly recognize a hierarchical bond between the participants. 573

Ultimately, both offender and victim negotiate the meaning of the coerced apology. That is, they decide whether it is designed to exclude or to include while humiliating. Jane must implicitly (or, if she is very sensitive to such matters, explicitly) decide if the inclusive-but-humiliating apology is something that she's willing to put up with. What's more, she should be clear, at least to herself, what acquiescence means. 574 Is it a recognition that the authority figure or group as a whole is just plain right? Is it a case where she wants to be part of the group so much that she's willing to pay her dues and play by the rules? Is it a case where she is being

---

573 Coercion always implies at least a provisional hierarchy, but the hierarchy need not be complex or even vicious. In Cinderella's case, she is being forced to debase herself in relation to her sisters and her stepmother. The sisters are also subordinate to their stepmother but may find this arrangement to be perfectly legitimate. Contrast this with a case in which a well-meaning stepmother coerces her stepchild and biological child into a forced apology while taking pains to structure the apology to minimize or eliminate feelings of humiliation, possibly by use of some appeal to universal principles and that "we're all on the same team."

574 That said, we often come to accept an inferior position gradually, without being aware that we are doing so. Abusers who "groom" their victims are experts at this kind of manipulation. Few people set out to take the weaker role in an abusive relationship.
treated *unfairly*? Or is it a case where she is upset, not necessarily because of a lack of fairness per se, but because she is not being treated the way she sincerely believes she ought to be treated? Her answer will contribute to her orientation toward the group and, ultimately, her determination as to whether or not the group's expectations and orders (implicit or otherwise) deserve respect and obedience.

If Jane is an insider, she will likely interpret the coercion as an appropriate if undesired exercise of legitimate authority. To use an analogy, a basketball player may loudly complain when the referee blows a call, but she will usually accept that referees have the authority to make bad calls. That said, unexpected coercion or coercion that is unjustified by established norms or sound reasons *that Jane can accept* – not that the group feels she *should* accept – may put her down the path of becoming an outsider. Such unexpected cases force Jane to re-evaluate her relationship with other group members and decide whether or not she is still part of the group in a meaningful, "political" sense.575

To continue the basketball analogy, let's assume that I join a pick-up game at my local park. Over the course of the game, other players complain that I am committing many fouls, but as far as I can tell, their complaints are random. Or, worse yet, I get the feeling that their complaints are stupid, wrong, or biased. As best as I can tell, their frustration and complaints are sincere. At this point I must reconsider my place is within the group. I begin to question whether or not we're playing the same

575 I want to stress yet again that becoming an outsider does not mean being free of all moral constraints with respect to the members of one's former group. If I decide to emigrate from the US due to what I take to be a drastic and unfortunate change in American political and social life, it does not follow that I do as I please to Americans who pass by my new home in Paraguay.
game, or if I am being treated fairly within the context of the game.\textsuperscript{576} I may complain, appeal to higher principles, protest, and so on, but at some I must decide that I am fundamentally alienated from the group, even if I am observationally participating the same way as my teammates.

Now let us return to Sue, the sibling who \textit{receives} the coerced apology. If Sue is an "insider", then she is in an analogous position to "Jane as insider." She may see her coerced acceptance as unfortunate but legitimate. Again, she need not accept the apology the way Tavuchis would prefer, regardless of whether or not a sincere apology were proffered. As with an externally motivated apology, she may have any number of reasons (which are not mutually exclusive) to accept the apology including, among others, a fear of reprimand by the authority figure, a fear of appearing unreasonable, a desire to demonstrate her allegiance to the social structure, or a desire to appear magnanimous in front of an audience – or to herself.

If Sue is an outsider, our assessment will depend on whether or not others recognize her as such. If she is clearly not a member of a given group, a coerced apology must be a universally recognized act of humiliation.\textsuperscript{577} This is because the coercer does not have recognized authority, even among members of her in-group, to demand that Sue perform symbolic acts that she would have to perform if she were part of the group. In other words, the coercer would be infantilizing Sue by presuming

\textsuperscript{576} "Fair" does not necessarily mean "equal." Perhaps it is appropriate that those with seniority are given additional consideration, for example.

\textsuperscript{577} I should specify that I am only talking about coercion between adults in this case. I acknowledge that there are cases in which an adult coerces an apology between children for the sake of teaching them socially acceptable behaviors.
an authoritative position. It is one thing to, when in a foreign culture, have one member of that culture exhort someone to apologize to you for an "offense" that wasn't offensive to you at all. It's another to have that person then insist that you play the role of "gracious recipient" of the apology.

If on the other hand, Sue is an outsider – perhaps she finds the group's values and structure reprehensible – but the group members believe that she is a meaningful part of the group, they may understand the coercion to be a legitimate case of an authority making a demand of a group member. A misunderstanding will thus take place: the group members will believe that the apology is establishing or reestablishing a particular type of relationship between secure group members, not instituting a humiliating relationship with an outsider. Yet their sincere but mistaken belief about Sue's relationship to them does not invalidate Sue's belief that the apology is a moment in which she is become more, not less, distant from the group. The apology may serve to highlight and strengthen group norms of reciprocity and respect, and these norms may be inherently unobjectionable, but it also further alienates some members who are drafted into the performance of the ritual.

For simplicity's sake, my example involved direct coercion applied by a parent to children. That coercive forces likely apply in deliberative settings. Coercion abounds in our personal relationships and relationships with close in-group members –

---

578 To be clear, the group may believe that the coercer has the right to make demands of outsiders but, in the case of apology, even they would recognize that the outsider is entitled to believing that the demand is illegitimate and that they are being coerced.

579 This is not the only possible interpretation of the event. She may believe that (1) the interaction is humiliating but that she is a real (if low-ranking) member of the group, or (2) the demand for an apology is a (possibly unwitting) invitation to become a real group member.
we are constantly under pressure to avoid disrupting how things are done. When we use more contrived fora, we face the unpleasant choice of either (1) accepting a possible free-for-all where socially powerful and confident members run roughshod over the weak, (2) a system in which a moderator "encourages" conflict resolution – usually in full view of other participants – by appealing to reasonableness, or (3) a rigid discursive structure designed to prevent members from actually engaging with each other, thus defaulting back to orthodox deliberation at best, or a mockery of deliberation at worst.

IV.2. The burden of acceptance

The problem of coerced apologies introduces another question that is at the core of the relationship between apologies, deliberation, and the philosophical anarchist critique: when an apology is made, who gets to decide whether or not it has been accepted? What do we make of a case when two people grudgingly apologize, shake hands, and turn their heads to sneer sarcastically at those watching the spectacle?

There are many reasons to accept an apology. In ideal cases, all parties are aware of the purpose of the apology and the apology is accepted in a way that is satisfying to the giver, the receiver, and the audience. Apologies, as I have shown, can be coerced and remain apologies nonetheless. They can also be accepted due to misunderstandings. I may be shocked and hurt when I find out that your apology was

580 I discuss this term earlier in the dissertation.
not a sincere expression of remorse but instead was coerced by an authority, and this would complicate my previously straightforward acceptance. I would want to rescind my apology, but somehow I feel trapped, especially if the apology was done in public. Apologies might also be accepted as ways of receiving social advantages and advancing one's social position; in some cases accepting an apology (in the right way) demonstrates magnanimity, maturity, and/or spiritual growth.

Yet apologies can also be accepted in ways that constitute real sacrifices, and it is in these cases that questions of coercion, real or imagined, become especially problematic. So far, my discussion has focused on people who want to fight to improve or at least maintain their social position, or else who may alienate themselves from the group. Many people, however are victimized from the start. People on the wrong end of power asymmetries routinely accept apologies in order to smooth things over or to help the group move on. This is not because they see the social order as desirable, but because they do not have the strength or social standing to take responsibility for the disruption of good order. They do not have the capital to be rude.

They – we – also routinely accept apologies because the alternative would be to throw a spotlight on painful, humiliating experiences or truths. Crucially, we may accept them because not accepting them would expose us as weak. The extreme version of this is to deny that an offense took place or that we didn't even notice the

---

581 Again, this is because to the extent that an apology is public, it becomes on-the-record speech.
582 That said, some people in leadership positions choose this strategy, judging it to be more effective than litigating potentially disruptive points. See for instance Willink and Babin, Extreme Ownership.
583 Consider the following sincere apology: “I'm sorry I successfully seduced your last four lovers.” The sincerity adds to the potential humiliation.
harm. We accept the apology, or try to say that no apology is necessary, in order to "prove" to others that we are not as pathetic as we believe ourselves to be.

This denial that harm was done adds lots of complexity to the question of legitimacy. On the one hand, the person goes on record as not having been harmed, or of forgiving the person who has harmed her. From the perspective earnest participants, what more can be asked for? On the other hand, this person would not have gone on record had social incentives differed. This is not a case of withholding information in order to maintain a sense of privacy or even for private gain. By denying the existence of the offense, the purported victim is outright lying, but may be justified in doing so.

This problem appears insoluble. The person who is deeply offended and receives an apology must relive the humiliating experience – in public – to fully accept it, knowing that doing entails the risk of being further judged or humiliated. After all, by showing that she was hurt over what others may consider a trivial offense, she may be accused of being too thin-skinned. This is no mere academic concern – people who have been routinely disrespected (say, for reasons pertaining to race) may have good reason to be hurt or angered by a "small" slight. In principle, this could be a learning moment where the offended person takes the time to recount just why they are hurt, but that is just too much to ask of a deliberative process, formal or informal. It is unfair to demand that the victimized person double down on his vulnerability in front of what he believes is a hostile crowd.

The upshot is that the person may want to reject the apology for many possible

584 Lazare, On Apology, 234.
reasons. He may feel that he was disrespected and that he has no duty to subject himself to further humiliation by engaging in a process where he negotiates the receipt of an apology. He may also believe that accepting the apology and "making things right" would violate his sense of self-respect, and that it would be wrong to do so. The apology, in other words, does not satisfy her needs in such a way that she can imagine herself as being a legitimate member of the group in a way that is consistent with her sense of self-respect.

What might this mean in deliberative contexts? In contrived deliberative settings, what do we make of someone who accepts a proffered apology but does so for prudential reasons? When the group comes to some decision, in what way is this person – who still feels belittled – bound? Two points come to mind. First, as I have said, we cannot resort to saying that she should have halted the proceedings and been clear about her feelings. Members of consistently disparaged groups, even if they have not internalized oppressive stereotypes, have very powerful incentives not to speak up. This has been dubbed a "social reality constraint," which "... makes it more difficult for [members of] the group to (publicly) claim parity or superiority." It would be admirable but supererogatory for a group member to speak up if she believed that she faced a hostile environment. Note that whether or not the other participants believe that the environment is hostile is not the issue. What is at issue is whether or not the offended believes it is.

Nor can we say that the solution is to create a supportive environment where

---

585 Spears, Jetten, and Doosje, “Ingroup Bias,” 341.
people could express themselves without fear, even in cases where there is no systemic oppression. Even we could create such a self-containted setting – a prospect which, given my remarks on etiquette in the previous chapter, I doubt is possible – we cannot unspeak painful truths once the session has ended.\textsuperscript{586} Exposing vulnerability to others and oneself may have salutary effects in the moment but end up being embarrassing once the group members get on with their lives. The fact that the group members are (usually) strangers does not offer much solace – some of our longest-held humiliating experience are those that took place in front of a nondescript crowd.\textsuperscript{587}

What's more, I have already argued via Moon in my chapter on deliberation that requirements for complete disclosure violate our capacity to shape our personalities, including our considered opinions, which are required for proper deliberation.\textsuperscript{588} In other words, we should not demand that people (1) be able to precisely articulate their feelings or (2) feel as though they have to work through these feelings with the very people who have caused the offense in the first place. We do not (or should not) deliberate by engaging in a stream of consciousness exercise designed to "get it all out there." We think, evaluate, shape our thoughts and then speak our mind. We will not do this perfectly, nor should we demand that others have a constant vote on what aspects of ourselves we feel are appropriate to share.

\textsuperscript{586} A point some romantic couples discover when intimate details, expressed in a state of exultation, come to weigh heavily on the couple's identity over time. Sharing intimate details may create a powerful bond, but it may also lead to estrangement and awkwardness even in what was a supportive relationship.

\textsuperscript{587} For instance, some of my most cringe-worthy memories involved incompetent attempts at seduction with people that I had good reason to believe I wouldn't see again. The anonymous nightclub crowds have terrified many a would-be dancer.

\textsuperscript{588} Moon, \textit{Constructing Community}, 91.
But second, the offense and the problems inherent in apology do not necessarily void agreements that came out of the deliberation. What do we make of the person who believes she has good reason to believe she has been offended and that the apology was insufficient to rectify the offense? In some cases, it means that she is not part of the "we." Perhaps she never was, or perhaps the offense was so grave or was so intolerable given her history of suffering social humiliations, that her status as an outsider came to be as a result of that particular offense. In either case, all this means is that she is outside the jurisdiction of one subset of the moral community – the in-group – while remaining within the jurisdiction of the broader moral community, one that dictates what we owe each other qua human beings.

I do not want to try to describe a full list of what we owe each other, nor is it necessary to do so here. If modern philosophical anarchists object to the state’s claims of sovereignty on the grounds that accepting those claims would violate their moral autonomy, then whatever justification they use to defend and articulate that autonomy will imply basic duties toward others. In the case of the offended person, getting rejected from a moral subcommunity, painful as it may be, does not seem to have any bearing on the duties that logically precede the existence of that subcommunity.\(^\text{589}\) For example, we can in the ordinary course of our lives come to agreements with people that we find profoundly disrespectful. These agreements are usually morally binding. The disrespect stemming from failed apologies may void the binding nature of the

\(^{589}\) Actually, this need not be true. It would be interesting to consider how rejection from a cohesive moral group may cause me to have fewer (or more) duties toward them then I would have had if I were thickheaded and interacted with the group members with no awareness that they formed a subcommunity at all.
deliberative procedures – the argument that "we all came to this determination together" in an appropriate way – but leaves untouched other types of moral glue.

Thus, offended philosophical anarchists may, and should, believe that they have duties despite their alienation from the group. Perhaps the offended person believes that she should nevertheless go along with the group decision for other moral reasons. For example, while she may dislike the people who have symbolically rejected her, she may believe that not liking people is an inadequate reason for trying to stymie them. Or perhaps the offended person does believe that some sort of restricted community exists, the kind that is restricted to one issue. This will turn on her broader substantive moral outlook, whatever that may be. Some people may believe that morality works that way, others not. Perhaps she believes (quite plausibly, even from a philosophical anarchist perspective) that agreeing to undertake a joint project creates not merely a duty to perform certain actions, but some important moral subcommunity, replete with unstated expectations, wants, hopes, and shifts in identity. This might, in turn, imply a duty to "go along" with some tangential demands of the group.

What's more, I want to emphasize once again that a person's beliefs, both with respect to what is morally correct as well as what is required for her to be treated with respect, may be problematic or even wrong. In cases of legitimate disagreement, there may be times when the side with more force will coerce the other in the name of a higher-order principle that the other rejects, correctly or incorrectly. My position is quite circumspect: I am saying that such cases should be seen as instances of coercion,
perhaps justified, but coercion of someone who does not take part in a meaningful political community.

Those who are doing the coercing may be mistaken when they believe that they are engaging with a fellow member of a political community. This is not so shocking – we are at times mistaken when it comes to what we thought were even our most intimate friendships. It is not implausible that we would make the same type of mistake when thinking about who does and does not take part in our meaningful political community.

V. Conclusion

In this chapter, I argued that apologies are important, that they are undertheorized in the deliberative democracy literature, that they are unavoidable in deliberative settings, and that they can easily go awry. Even when "successful," they can cause proceedings to fail to integrate all deliberators. The result of these problematic apologies is that either someone will have good reason to believe that her sense of self-respect has been violated and/or that the group will act in a way that violates the normative tenets of deliberative democracy. As far as this dissertation is concerned, this amounts to the same thing: the person in question, while retaining the general moral duties she has to all people, is excluded from the political community, which is a subset of that broader moral community.

This conclusion is easy to miss because apologies frequently appear to be at least superficially successful. Actual apologies implicitly refer to an idealized, dyadic
model in which the participants behave in predictable ways – call, apology, response. Yet actual apologies, while being observationally equivalent to ideal apologies, oftentimes serve very different social purposes. This observational equivalence effectively silences vulnerable members, or members who feel slighted, much the same way that an exquisitely crafted polite but cutting remark can leave us without any response that doesn't make us look even more foolish. An apology was made, but something nonetheless feels very wrong.

In sum, respectful deliberation requires but can be undermined by apologies and etiquette. Trying to determine ahead of time what counts as a “respectful” apology is as futile as trying to pin down the meaning of a word without seeing how it is actually used. What’s more, in an increasingly diverse polity, we do not have the luxury of relying on deeply shared meanings and practices to adjudicate between “correct” or “incorrect” demonstrations of respect in everyday encounters. Even if we did, mere correctness does not imply that the “correct” response – such as lowering one’s eyes when talking to a social superior – is right. Nor can we always say that a response is wrong. When non-malicious citizens come together to discharge their duty to try to deliberate, each must decide what it would take for them to have to end the deliberation in order to maintain their self-respect. Philosophical anarchists who try to deliberate may come to find that they have political obligations after all. Or they may not. The only way to know is to go into the world find out.
Chapter 9: Rehabilitating Wolff

I. Introduction to the Conclusion

When I began researching the philosophical anarchist challenge to political legitimacy, I expected to settle on and endorse a vigorous defense or rebuttal. Years later, I am forced to disappoint those few friends and colleagues who want to know “who wins” the debate. Far from coming up with an unambiguous answer, I have argued throughout this dissertation that no theory of obligation decisively defeats the philosophical anarchists on their terms.

That said, some theories work sometimes. I have argued that natural duty theories were particularly likely to persuade, especially when used indirectly. In my chapter on consent, I noted but did not spend much time investigating the viability of a theory that posited a duty to vote, with consent theory then finishing the job (via the symbolic import of voting) of binding philosophical anarchists to the state. Later, I argued that there is a duty to attempt to deliberate, but that this deliberation can easily fail to legitimize, allowing philosophical anarchists to claim to have discharged their natural duties without having to endorse the legitimacy of the state.

There is one more radical attempt to come to a decisive answer to the philosophical anarchist question, and that is to side with Wolff over Simmons. In other words, we can try to argue, alongside Wolff, that the legitimate state is an impossibility. After all, nearly this entire dissertation has considered a posteriori philosophical anarchism, something that Wolff would consider too great a concession. If we commit to this strategy, we would have to defend a very unorthodox position – a
priori philosophical anarchism – within an already unorthodox field – philosophical anarchism writ large. And while a full defense of Wolff would constitute a dissertation-length project on its own, in this final chapter I want to argue that Wolff’s position is more plausible than its reputation implies.

As I said in Chapter 1, many commentators have attacked Wolff on various grounds. Rather than try to provide a general defense of Wolff, I will dispute what I think is the critics’ most serious attack: that Wolff’s position precludes him from being able to endorse simple acts of promising. The critics who levy this charge are using a *reductio ad absurdum* argument here. Surely Wolff does not want to say that simple promises are impossible or a violation of moral autonomy, but these critics argue, that this is precisely what his framework entails. If this is correct, Wolff’s project – and probably a priori philosophical anarchism writ large – is truly dead in the water.

The crux of the charge is as follows. First, Wolff argues that accepting political authority involves an abdication of some or all of our autonomy. Wolff also believes that our primary moral duty is to act according to our autonomous dictates whenever possible. Our commitment to acting autonomously implies a duty to safeguard this autonomy. Acknowledging an obedience to obey the state implies an agreement to act contrary to one’s autonomy. Thus, Wolff argues, we cannot morally accept the existence of legitimate political authority.

But Wolff’s critics argue that his radical position vis-à-vis state authority should lead him to reject the institution of promising as well. After all, promising to

---

590 This rehabilitation has occurred in fits and spurs.
obey someone – even in a limited capacity – is ostensibly an abdication of one's autonomy. When we promise, we commit to performing an action even if, when it comes time to execute that action, we think it would be right (or otherwise preferable) to do otherwise. We thus bind ourselves to the performance of an action or – worse yet – to the dictates of another. Why should this differ from a promise to obey the state?

This is not the first time that a philosophical anarchist has been accused of being incapable of endorsing the institution of promising. Critics such as Carole Pateman have argued that Wolff should look to William Godwin, an 18th-century philosophical anarchist who embraced the position that we should ultimately reject promising. Yet this suggestion does not do justice to the breadth of philosophical anarchism. Philosophical anarchism is a large tent comprised of among others, Christian anarchists such as Tolstoy, proto-Nietzscheans like Max Stirner, and modern-day consent theorists such as Simmons. Philosophical anarchist positions are not interchangeable – Wolff is not Godwin. Godwin’s argument against promising amounts to an act-utilitarian objection to the binding nature of promises. He believes that we must be able to change our plans at any given time; to bind myself now means to possibly forgo selecting the best option in the future. Wolff, on the

593 When we promise, he believes, “[...] we abridge, and that is the most essential point, the time of gaining information, if we bind ourselves today to the conduct we will observe two months hence. He who thus anticipates upon the stores of knowledge, is certainly not less improvident than he who lives by anticipating the stores of fortune” Godwin, *Concerning Political Justice*, Book III, chapter iii. I should note that Godwin’s position amounts to more than a crude utilitarianism based solely on rational calculation. He also subscribes to a perfectionist view similar to that of J.S. Mill. He believes that when it comes to our capacity to think and make decisions, we need to use it or lose it.
other hand, considers himself a (not wholehearted) follower of Kant. If Wolff ultimately rejects promising, it will have something to do with autonomy, not with Godwin's epistemic concerns or with appeals to utility (or something that looks like utility).

This chapter addresses two versions of the “Wolff cannot accept promises” challenge. I have called these challenges the “Alienation Critique” and the “Promises are Promises Critique.” The Alienation Critique, put forth by Graham Baugh, levies the following charges against Wolff. First, it holds that Wolff's conception of autonomy precludes us from developing social conventions such as promising. What's more, he believes that even if we concede that Wolffians could develop some social conventions, Baugh believes that promising would not be one of them because all promises in a world of Wolffian agents would amount to cheap talk. This is because, on Baugh's interpretation of Wolff, all Wolffians would preserve their autonomy in order to pursue their own moral goals. Promises would be broken the moment they came into conflict with these goals. I argue that this interpretation is simply inconsistent with crucial elements of Wolff's moral theory. Although Wolff explicitly distances himself from Kant, the Alienation Critique unfairly exaggerates this distance.

The Promises are Promises Critique, expressed by (among others) Simmons, states that the very reasons that lead Wolff to argue against the validity of promises to obey the state also apply to promises writ large. What's good for the goose must be

595 I address his divergence from Kant in the next section of this paper.
good for the gander – if my promise to obey the state violates my autonomy, so must my promise to meet my friend for a drink this afternoon. In other words, those who argue this position believe that Wolff’s critique applies to a class of actions which contains all instances of promising. I will argue that this critique conflates different types of promises. Although we may say 'I promise' in two cases, these promises may amount to very different types of symbolic action. Even though Wolff does not provide an explicit response to the Promises are Promises Critique, his work contains resources that would allow him to differentiate between types of promises. Some but not all promises will entail a violation of autonomy as he understands the term.

II. The Alienation Critique

II.1 Summary and description of Wolff

Baugh puts forth what I call the Alienation Critique (AC) in his 1985 essay, *The Poverty of Autonomy*, where he argues that Wolff cannot provide a coherent account of promising for two interrelated reasons. First, Baugh believes that Wolff’s understanding of rational agency is so individualistic that it cannot account for the existence of intersubjective meanings and conventions. Promising is one such convention, so, if Baugh is correct, Wolff’s autonomous individual either lives in a world without promising or cannot fully understand and participate in any convention of promising that does exist.

Second, Baugh believes that Wolff’s autonomous person is characterized by a relentlessly instrumental rationality, one that pushes the individual to pursue her
chosen goals as efficiently as possible. If this is right, then any promise would become void when the promisor either comes to have a different goal or when she decides that she can best achieve her goal by reneging on the promise. So even if we reject the first claim – that Wolff's conception of rationality is too solipsistic to account for intersubjective meanings – Baugh will say that Wolff's ethical commitments still lead him to the same conclusion, namely that promises are simply not possible in Wolff's framework.

Baugh's objections, which I will address more fully below, fundamentally turn on his understanding of Wolff's critique and partial appropriation of Kant. In the preface to *In Defense of Anarchism*, Wolff says that he has "simply taken for granted an entire ethical theory." This statement, taken alongside his uncompromising insistence on the primacy of moral autonomy and his explicit reference to Kant, leads us to believe that he something very close to a Kantian deontologist.

Wolff fleshes out his position in his subsequent work, *The Autonomy of Reason*, where he offers an extended and ambivalent assessment of Kant's moral philosophy. Kantian freedom as Wolff understands it means not having one's behavior determined by heteronomous, contingent forces. Rather than acting according to these forces, the free person's actions are determined by moral law. This moral law is rational law. There is for Wolff no real “freedom” in the colloquial

597 Wolff, *Autonomy of Reason*. Incidentally, many of Wolff's critics make no reference to this work, even though it provides crucial details and in some cases revisions to Wolff's position in *In Defense of Anarchism*.
598 In spite of his sympathy with the struggle of marginalized groups, Wolff uses the term “free man” throughout his text. I replace “man” with “person” throughout this paper.
sense of the term – either we are obeying moral law or we are effectively getting pushed around by other forces.\textsuperscript{599} This is not necessarily depressing since Wolff agrees with Kant that we cannot know that we are not free (in some sense) in the noumenal realm.\textsuperscript{600} He further agrees that we can presuppose that we are free in that sense, even if we can't say what exactly that kind of freedom amounts to.\textsuperscript{601}

Wolff and Kant agree that while we cannot intuitively determine whether or not our aims correspond to rational law, we have recourse to the categorical imperative (CI) test as a mechanism to verify whether or not a potential action (or instance of inaction) is rational. But here Wolff parts ways with Kant. While Wolff agrees that the CI test rules out actions that are impermissible, he disagrees with Kant's assertion that the CI test can generate \textit{positive} moral law.\textsuperscript{602} He wants to retain Kant's demand for consistency while rejecting the position that rationality requires adhering to more substantive positions.\textsuperscript{603} He is unyielding on this latter point: “The real truth of the matter, I think, is that Kant fails to find a plausible argument for the validity of substantive moral principles because there simply are no such principles!”\textsuperscript{604} Yet he immediately follows this statement by declaring himself to be both a 'follower of Kant' as well as a consent theorist:

\textsuperscript{599} Wolff, \textit{Autonomy of Reason}, 216.
\textsuperscript{600} Ibid., 222.
\textsuperscript{601} Ibid., 217.
\textsuperscript{602} The critique that Kant's ethical theory cannot provide adequate substantive principles is of course well established, and has been since at least Hegel's \textit{Philosophy of Right}. Wolff makes no mention of this. This is unsurprising when one considers Wolff's longstanding anti-Hegel stance. From his “mini-tutorial” of Marcuse's \textit{One Dimensional Man}: “Faithful readers will know that I have an allergic reaction to Hegel, so I consider it an evidence of my admirable broadmindedness that I am willing to take Herbert's works seriously, as I do.” Wolff, “One-Dimensional Man”.
\textsuperscript{603} Wolff, \textit{Autonomy of Reason}, 217.
\textsuperscript{604} Ibid., 213.
I am persuaded that moral obligations, strictly so-called, arise from freely chosen contractual commitments between or among rational agents who have entered into some continuing and organized interaction with one another. Where such contractual commitments do not exist, cannot plausibly be construed as having been tacitly entered into, and cannot even be supposed to be the sort that would be entered into if the persons were to attempt some collective agreement, then no moral obligations bind one person to another.605

Wolff's deontology is thus much thinner than Kant's. Wolff believes that all social obligations are necessarily contractual, and this is why he objects to Kant's positive duties, e.g., Kant's imperfect duty of beneficence. Wolff believes that these positive duties are not required by reason; they are the kind of duties that one may choose to adopt, not duties that we must accept.606 But rather than try to provide substance where Kant supposedly fails to do so, Wolff simply says that positive duties to others only emerge as a result of contracting and thus cannot be specified in advance. The upshot is that while Wolff has much to say about what we ought to avoid, he leaves us without instruction as to what we ought to do other than to keep our word. Baugh seizes on this lack and argues that it leads to an overly individualistic understanding of morality. Kant's positive duties, the story goes, bind us to a common substantive moral framework in a way that Wolff's moral theory does not.

605 Ibid., 219. Readers interested in consent theory should note that Wolff's last sentence blurs the line between different types of consent. He begins by saying that there are no binding obligations if people have not actually entered into an agreement with each other. By the end of the sentence, he is referring to what Pitkin coined hypothetical consent—a pseudo-consent purportedly justified by a belief that the agent would have consented had she been fully informed, had access to her intellectual faculties, etc. Pitkin, “Obligation and Consent I”; and “Obligation and Consent II.”

606 Take for instance Wolff's commentary on the purported duty to cultivate one's talents: “Nothing much need be said about this argument for cultivating one's talents. Considering the oppressive odor of relentless moralism which clings to the passage, it is just as well that the argument is so obviously bad”. Wolff, Autonomy of Reason, 169.
II.2 Baugh's critique

In this subsection, I focus specifically on reconstructing the portions of Baugh's argument that support his position that Wolff cannot make promises. To begin, Baugh worries that Wolff understands rationality in purely instrumental terms. Baugh believes that Wolff thinks that autonomy is the only intrinsically morally relevant goal an agent may have and that “Any agent who acts according to self-chosen policies, which are internally consistent and rationally connected to her goals, is autonomous in a Wolffian sense.”607 (Baugh 2002, 115). In other words, Baugh interprets Wolff as saying that it doesn't matter what I do as long as my actions help me achieve whatever goal I happen to have. The fact that my goal may harm others is ostensibly irrelevant from this standpoint of Wolffian autonomy – a sensible statement if it's true that Wolff is only concerned with non-contradiction. But does this mean that discharging this duty to oneself will lead a Wolffian to engage in asocial or antisocial behavior? Baugh thinks so, and that this cannot be helped if the duty to preserve one's autonomy has lexical priority over other duties and if the duties that each member of a community holds – duties which follow from idiosyncratic goals – end up in conflict.

Baugh's worries about Wolff's conception of autonomy go even further. Baugh believes that Wolffian autonomy ironically divests agents of all responsibility. The idea is as follows. We can presumably agree that we should not be held responsible for nonrational behaviors – our responsibility for an action is eliminated or at least severely attenuated if the action is the result of, say, a tic. But, Baugh says, since

Wolff believes that we do not choose our goals or values – in other words, our conception of the good – we cannot (under a Wolffian framework) be held responsible for having them. Baugh understands Wolff to be saying that we can and must choose to pursue those goals as efficiently as possible. We may be responsible for selecting subordinate goals that we believe will help us actualize the more fundamental goals, but if the fundamental goals and values are unchosen, we remain fundamentally unfree. Worse yet, these unchosen values and priorities may or may not take others into consideration.

If Baugh is right that Wolff believes that our primary duty to is act autonomously, it seems that we cannot blame someone who acts instrumentally rationally in the service of her fundamental goals or values. We may oppose her when her fundamental goals conflict with ours, but we cannot blame her. In fact, we should in some meaningful sense praise her for acting autonomously and not kowtowing to our conception of the good.

Still, we may be disturbed by the thought that the very best we can hope for is either a miraculous convergence of 'given' goals or a tense modus vivendi. Baugh seems to believe that a promise in such a world would have little practical value and that I would renege on any promise that I made once it is to my advantage to do so. If my goals are unchosen, they can presumably change without any input on my part. Even if my fundamental goals do not change, perhaps changing circumstances (or even further reflection on my part) will lead me to believe that I should undertake a

---

608 Ibid., 116.
new course of action that would render my promise void.

Not that we would even get that far. Baugh believes that Wolff's 'extreme' individualist conception of rational agency prevents agents from even communicating intelligibly with each other. As he says,

Wolff cannot provide a coherent account of obligation either, for obligation arises from various social practices, such as promising, and Wolff's purely individualistic and abstract analysis of rational agency simply cannot provide a coherent account of any social practice, for social practices are constituted by intersubjective meanings and impersonal rules which transcend pure, subjective rational agency and arbitrary choice.\(^{609}\)

This is Baugh's most radical critique. If he is right, then a Wolffian would not even get the chance to break her promises willy-nilly, since she would not even be able to understand what a promise \textit{is}. In other words, Baugh believes that Wolff's conception of autonomy presupposes that we each have a mental model that is 'given' and unable to adapt to the mental models of others, except insofar as it can take external stimuli and process them from its own perspective. Baugh seems to be saying that a Wolffian would say that our mental models differ so much that what might seem like conventions would lack sufficiently shared meaning to make them worthy of the name.

It is not clear just how far Baugh takes this line of thinking – does he believe that Wolff's position entails that there is no possible \textit{agreement}? Or does he believe that Wolff's position amounts to something close to Wittgenstein's famous statement that “If a lion could speak, we could not understand him.”\(^{610}\) Whichever interpretation

\(^{609}\) Ibid., 117.
Baugh favors, we come to the same conclusion: if Baugh is correct, Wolff cannot endorse promising.

While this does not exhaust Baugh's full critique of Wolff, it shows why Baugh believes that Wolff's position precludes promising. To re-iterate and summarize: Baugh believes that moral discussion is impossible in a world of Wolffian agents. Wolff's apparently eccentric commitment to an odd form of autonomy – the efficient pursuit of unchosen goals – leads to an overly individualistic understanding of morality. Any promise becomes cheap talk since one's fundamental goals or the best strategies to achieve one's goals could change without notice. Worse yet, because Wolff's agents (as Baugh understands them) may have radically different goals – both moral and practical – they may end up being unable to communicate intelligibly with each other.

Both of these interpretations fail to describe Wolff’s position adequately. In the following subsection, I hope to show that Baugh is unfairly imputing to Wolff an overly bizarre theory of mind. Wolff's autonomy is far more banal than Baugh makes it out to be, and this saves him from Baugh's critique.

II.3 Response to Baugh

In this subsection, I respond to the preceding objections in an expansive manner. I hope to not only respond to Baugh specifically but also to use my response to him as a platform to deal with other potential objections.

Baugh's critique is ultimately unpersuasive. Most obviously, Baugh is mistaken
when he says that Wolff's autonomous person would break a promise if doing so were instrumentally required to pursue another goal. This mistake stems from his unfair critique of Wolff's use of the CI test. Recall that for Wolff, adherence to the CI test does not generate universal positive duties. Baugh believes that this would lead people to hold incompatible goals, with each person morally obligated to try to actualize her own goals.

Yet Baugh does not seem to appreciate that the test nonetheless does effectively stop a follower of Wolff from breaking promises. Even if the autonomous person's goals are, so far as we can tell, completely arbitrary, the autonomous person judges the acceptability of her own goals by running them through the CI procedure and may do so because the CI test ensures that the person acts according to the principle of non-contradiction.

According to Wolff, any goal that satisfies the CI procedure prior to any promises being made is acceptable, but the pursuit of a goal would be inadmissible if doing so would lead to the violation of a promise. To use a banal example: it's fine for me to spend next Monday napping or meeting with students in order to give them extra preparation for an upcoming exam, but once I agree to meet with students, I cannot unilaterally renege and decide to nap instead. My decision to meet the students is not an indictment of napping in general; it just means that I can't nap on that particular day.

Perhaps my unilateral decision to nap would be defensible if I realized (after making the promise to my students) that my napping was necessary to satisfy some
prior commitment that I made, for instance, that I be well rested for a late night helping my daughter with her science project. But this is not the arbitrary reneging that worries Baugh. He is worried that I will renege simply because my attitude changed on Monday morning or that I decided that napping would be the right thing to do all things being equal. Yet this worry is unfounded – running the "napping maxim" through the categorical imperative test after making the promise would yield a contradiction that would not have existed had I ran an analogous "napping maxim" through the test prior to making the promise. Wolff's position is that my attitudes and goals inform the promises I make, but once a promise is made, my status as an autonomous agent obliges me to abide by the terms of the promise.

Baugh is mistaken when he says that Wolffians cannot make agreements. What he should be saying is that some agreements are not possible because they would predictably bind us to positions that would force us to violate our other moral obligations. Many of these obligations come about due to prior agreements. The only way for Baugh's critique to hold is if Wolff held the position that the values and goals of a given individual could change\textsuperscript{611} in such a way that the individual would find the CI-test silly or irrelevant. Such a person could break promises with impunity since she would no longer be bound by the non-contradiction clause of the CI test. Unfortunately for Baugh, Wolff is quite clear that this would not happen. Or, more precisely: if I experienced such a conversion and decided to ignore the CI-test, then my moral outlook would simply be wrong. The CI-test determines whether or not my

\textsuperscript{611} I am using the passive voice here quite deliberately. Hopefully, I have shown by now that Baugh doesn't think that a Wolffian agent could deliberately revise her fundamental goals and values.
actions conform with freedom; I cannot “freely” choose to discard it. Doing so would mean to act heteronomously, and thus fail to uphold my autonomy in a Wolffian sense.

The objection that Wolff’s autonomous person can switch goals arbitrarily is also odd. Why is it a problem if a follower of Wolff changes her mind on morally pertinent issues? In cases where the change is self-regarding, it is unproblematic. But what about cases that involve a change in how I believe I should treat others? Again, Wolff seems clear on this point: if someone decides that her goals were ill-advised – perhaps she has come to revise her theological convictions – she may pursue other goals so long as these goals pass the CI test and do not violate existing agreements.

In the extreme, perhaps this does involve reneging on some agreements. Yet even this is not necessarily a problem. We should be happy when former racists renge on their sincerely made commitments to try to foment a race war. What's more – and I will address this point in the following section of this chapter – some moral commitments cannot be kept or cannot be made. While there may indeed be tragic situations where someone is bound by prior commitments that she now finds unethical, part of being a conscientious moral agent is being careful when making commitments. Only those who fail to take their moral duties seriously when making agreements end up constantly facing these dilemmas.

What's more, Baugh's worry that an individual might suddenly shift her goals or values is overly theoretical. First, most people have remarkably stable – or at least

---

612 Indeed, it would be more accurate to say that the goals would switch on the person, seeing as how Baugh believes that Wolff that we have no choice in the matter.

613 Again, this provides support for Wolff's implicit position that some agreements are invalid from the outset.
slowly evolving – ways of understanding the world. Worrying about sudden value shifts in those around me is analogous to worrying that one of those people might suddenly snap and attack me. People do snap, of course, but fortunately those cases are rare enough for us to live our lives without seriously attending to that possibility.  

Second, even if an individual (1) suddenly shifted goals and values, and (2) were strictly instrumentally rational in the way that Baugh describes, she would nonetheless recognize the social value in maintaining consistent behavior. Indeed, if Baugh is right and Wolff's autonomous person is instrumentally rational, then this very fact demands that she act consistently in most cases. In other words, her formulation and implementation of an instrumentally rational plan must take into account the very large value in being seen as someone who acts according to consistent principles. We need others to accomplish our goals – especially our moral goals – and those who constantly renege on agreements will fail to get others to cooperate. Even the caricatured Wolffian that Baugh describes would, by the very fact that she is such a calculating machine, act consistently in order to get the support she needs to make her real goals come to fruition.

Finally, I admit that I do not fully understand Baugh's contention that Wolffian agents could not enter into agreements because they would be unintelligible to one another. Pluralism – both value and linguistic – does not prevent the development of

---

614 Those who habitually snap are, generally speaking, already understood to be "the type of person who snaps." Thus, while they might go from calm to aggressive quite quickly, we are less and less surprised each time they shift their demeanor.
shared language. Even if two people do not share *umwelts* and aren’t “really” communicating in a deep way, they can still communicate in such a way that their behaviors are mutually predictable. In fact, Baugh seems inconsistent on this point. On the one hand, he seems to think that Wolff’s position precludes social conventions of any kind, and on the other, he says that

> Any moral discourse between such agents will consist of competing but incommensurable and mutually unintelligible claims. In Wolff’s anarchist community there could be no meaningful debate or discussion of collective choice of ends. There could be a limited rational discussion of policies but only in an instrumental sense.  

Clearly, this type of discussion would not be possible if Wolff were as radical as Baugh says. But Wolff just isn’t all that radical; he is simply committed to a principle of consistency that prevents wanton promise-breaking. In fact, Wolff’s concern with non-contradiction would make him even more sensitive than most to the evils of going back on one’s word.

### III. The Promises are Promises Critique

#### III.1 Are all promises alike? Or, how to fail at promising.

The second critique is what I call the Promises are Promises Critique (PaPC). In formulating it, I draw heavily on Simmons' relatively brief remarks on the matter in his essay, *Philosophical Anarchism*. Simmons begins by pointing out that Wolff says that not only can a state not unilaterally declare itself to be legitimate, but it also

---

616 Simmons, “Philosophical Anarchism,” 36-37.
does not gain legitimacy even if we promise to obey its commands. According to Wolff, promises, in this case, do not bind because they would violate our duty to remain autonomous by subordinating our actions to the will of others. A promise to obey authority is possible; it's just not morally permissible.

Simmons complains that Wolff does not say why promises to the state are especially egregious to autonomy. After all, if I were to promise my friend to help her do something, my friend would have authority over me, at least to some extent. To promise is to authorize – to give authority to – someone or some group to demand action on our part. To expand on Simmons a bit, I want to add that open-ended promises are especially problematic. “I promise to let you decide what we do tomorrow evening” is both a common and formally terrifying promise; after all, how do I know you won't decide to forgo watching a movie in favor of robbing a bank or, more legal but no less repugnant, have us write angry letters to your ex-spouse? For someone like Simmons, Wolff can only be consistent if he rejects the legitimacy of any promises for the same reasons that he rejects state legitimacy. Simply put: if a promise to obey the state is invalid because it would violate our autonomy, what gives us the right to make any promises at all?

Simmons' challenge is in one sense quite odd. More precisely, it's odd that this type of challenge comes from Simmons. He is, after all, a consent theorist par excellence – the central theme running throughout his body of work is that consent is the only plausible basis for state legitimacy. Consent theorists have long

---

618 Ibid.
acknowledged that acts or utterances that are understood to signal consent may fail to do so, and Simmons is obviously aware of these restrictions. It is thus surprising that he would assume that a supporter of Wolff would think that all promises are created equal. They clearly are not. We can see this by considering promises that are stillborn at the moment of their conception.

Theorists differ on the particulars, but, as I described in a previous chapter, Keith Hyams has recently cataloged four ways that an act of consent – including a promise – may fail to “come off.” It may fail if an individual is generally too incompetent to give consent in any case, if the person does not understand that her words or actions actually signaled consent in a given case, if an individual's alternative options are sufficiently unpalatable, or because an individual is not morally authorized to give consent given the substance or scope of the attempted agreement. Before accusing Wolff inconsistency, we should investigate the possibility that his position amounts to saying that a promise to obey the state fails for one of Hyams' reasons.

If Wolff is right, then the promise to obey the state must fail because it violates Hyams' final criterion, as the other reasons clearly don't apply. Going down the list, (1) we can assume that most people are competent enough to give consent in a general sense. (2) An individual who says “I consent to obey the laws of this government”

---

619 As an example of this latter point, I cannot normally consent that you break into my neighbor's house and make off with his television set. For a discussion of all the criteria, see Hyams, “Consent Doesn't Work,” 112-19.

620 There is of course disagreement about the quality, scope, and limit of each of these restrictions. For instance, should it matter whether or not my options are limited as a result of coercion or mere bad luck? My point is not that these are settled issues, but rather that the debate is about the scope and not the existence of constraints on consent.
surely knows what those words mean.\textsuperscript{621} (3) A Humean objection that we have no real (practical) choice but to obey the law misses the point here. Wolff does not base his rejection of promising on the grounds that we have no real choice but to obey; he believes that such a promise would fail even if a person \textit{wanted} to make it.\textsuperscript{622} If promising in this case is objectionable to Wolff, it must have something to do with either the scope or the content of the promise.

We can clarify this further by taking up the argument of Wolff's most persistent critic, Jeffrey Reiman. Shortly after Wolff published \textit{In Defense of Anarchism}, Reiman responded with an extended rebuttal entitled \textit{In Defense of Political Philosophy}. Reiman argues (among other things) that Wolff objects to what Reiman calls a "wild-card morality," which Reiman represents with the symbol *. According to Reiman, Wolff objects to a promise to do * if someone else may determine the content of * in the future. Reiman agrees with Wolff that any such promise must fail and is inconsistent with the exercise of moral autonomy:

\begin{quote}
Since the specification of the source [of the content] does not specify the moral nature of the action to be substituted for the 'wild card,' \textit{such a morality entails a moral obligation to do an action with indifference to the moral nature of that action}, which is absurd.\textsuperscript{623}
\end{quote}

But, Reiman continues, political legitimacy comes not from a promise (and thus he agrees with Wolff on this point) but on a belief that we are better off with a system of

\textsuperscript{621} There are two points to mention here. First, the situation may not be as clear-cut as it might appear – I discuss the meaning of this type of consent later in this section. Second, I leave out cases of tacit consent, which are far more complex. In such cases, not only do we need to establish that the tacit action (or non-action) could constitute consent, but also that individuals are aware that performing this action (or non-action) constitutes consent.

\textsuperscript{622} Wolff, \textit{Defense of Anarchism}, 16-17. Like other failed promises, the point isn't that the promise is unintelligible, but rather that it is morally indefensible.

\textsuperscript{623} Reiman, \textit{Defense of Political Philosophy}, 3.
laws than without it. Reiman thus agrees with Wolff that we cannot give broad consent to obey the laws while also saying that Wolff's whole approach to political legitimacy is doomed from the start. And while Reiman's objection can and should be taken up elsewhere, his idea of “wild-card morality” helps Wolff here.

Reiman's belief that it is absurd to promise to * lines up with Hyams's final case of failed consent. In Hyams' account, a promise may fail due to its substance or scope in a variety of ways. For instance, a promise is invalid if the promise's jurisdiction requires prior authority that the promisor does not have. I can promise my students that they will pass my class if they turn in all their assignments and write sensibly, but I cannot promise that this minimal level of work will be sufficient to pass my colleague's class. To take another example, one injunction (e.g., Locke's) against suicide states that we do not have the authority to take our lives because doing so would be a violation against God. To those who hold those metaphysical beliefs, my promise to kill myself if I don't win a medal in the next Olympics would be void.

What's more, a promise may fail because we believe the content of the promise to be evil. By some – hopefully many – accounts, a promise to kill my dog if I lose a bet would be void because the substance of the promise requires me to violate a value that antecedes the promise. This would hold even if I lived in an area which legally

---

624 Ibid., 3-9. As Reiman unfolds his argument, he begins to look very much like a rule utilitarian. Indeed, Wolff assumes in his Response to Reiman that this is Reiman's position. Reiman, however, rejects this label. Reiman, “Anarchism and Nominalism,” 105.
625 Locke, Second Treatise, 9.
626 As with Hyams' restrictions, I do not want to imply that I think that the question of what counts as evil has been settled. I am pointing out that there exist some actions which do not violate appropriate jurisdictions but are nonetheless void because of the nature of the promised act.
allowed dog owners to kill their pets at will. Analogously, the problem with a promise to obey the state is that it is a wild-card promise that may require us to either violate the promise in the name of morality or to perform an unjustifiable action.

III.2 Attenuated conception political promises

A critic might agree that Wolff is in effect arguing that a promise to * is invalid because (1) its scope is too large, (2) it may commit someone to perform immoral acts, and (3) it directly violates the duty to maintain and act upon one's status as an autonomous being. Yet, this critic might continue, both Wolff and Reiman are mistaken when they say that a promise to obey the state means the government can fill in * with whatever it wants. As numerous commentators have pointed out, few defenders of state legitimacy seriously believe that this is what a promise to obey amounts to.\textsuperscript{627} Consent to obey the government is usually understood as a commitment to obey limited government, one that does not have an unlimited scope when determining *. So it seems that if Wolff objects to promises to * on the basis of *'s scope, the objection can be overcome if the promisor is the one who can set the limits of *'s scope. But a supporter of Wolff might reply that this does not obtain when promising to obey the state. A promise to obey a state is a promise to * or a promise to perform certain actions with the understanding that the state can change those actions without rendering the promise null and void. We don't just promise to obey the state as it is; we also implicitly promise the state as it will be.

\textsuperscript{627} Beauchamp and Witkowski, “Pure Anarchism.”
The upshot is that a promise made to a state is at least in this respect a fundamentally different undertaking than a promise to a friend. When I promise my friend that I will perform an action, I determine the boundaries of that action. When I promise the state to obey the law, I do so in a context where the state claims the right to modify the scope of the law. The state may in some cases be justified in making this modification, but this does not imply that I have a duty to agree to take part in such an arrangement, or even that consent to enter such a lopsided agreement should count as consent at all. More importantly, any agreement that takes place is no agreement between equals, or even between non-equals that have similar bargaining power. It is reasonable to refuse to consider any agreement legitimate, however substantively benign, when one participant is an all-powerful institution that has the de facto ability to defect or modify the terms of the deal at any time.

III.3 Promises and the social meaning of agreements

Those who defend the legitimacy of the state often use a particular dodge that tries to preserve the ghost of legitimacy while conceding nearly everything to those philosophical anarchists – those who superseded Wolff – who say that political obligation is possible but rare. These defenders extend the argument that we only consent to limited government by saying that we also consent to obey as long as there are no clear countervailing reasons that require disobedience. The fact that something

628 Assuming, as Godwin noted, that it even makes sense to say that I've consented to obey the giant and ever-changing corpus of law, the vast majority of which I will never get a chance to read. It might be more accurate to say that I consent to waive my effective right to consent to obey the law.
is illegal, they go on to say, means that there must be at least a *prima facie* to obey the law. The weight of this obligation is difficult if not impossible to specify, but it would be greater than nil and far below a trump. These defenders of legitimacy want philosophical anarchists to admit that the law carries at least a vanishingly small amount of weight. They want an admission that the law is supreme even if this admission has little or no practical bearing on actual moral decisions.

This position represents the ultimate rapprochement between the a posteriori philosophical anarchist and “archist” positions. Philosophical anarchists like Simmons say that adopting philosophical anarchism does not entail a change in one's behavior. Defenders of political legitimacy who concede that morality may require one to frequently disregard of the law are essentially saying the same thing but from the other side: the fact that something is *legal* might not have any bearing on any of our practical decisions. Something's being the law would have as much force as the Queen of England's opinion on a crisis in Canadian politics.

It is important to note that this is *not* what Wolff is proposing. Wolff's position is incompatible with such a rapprochement. There is a difference between breaking a promise that has weak obligatory force and deciding that a promise is void from the start. Wolff is saying that that the promise to obey is void, *not* that the promise is obligation-generating but overcome by weightier considerations. Wolff's position gives philosophical anarchism far more substance than does Simmons.

Finally, Wolff's critics may retort that the fact that most people understand

---

629 Smith, “Prima Facie.”
630 Simmons, *Philosophical Anarchism*, 20.
“She promised to obey the law” as being both sensical and duty-generating shows that
the promise does is not inherently void. To make such a promise, in other words, is not
to utter an absurdity. On the contrary, these critics may continue, such a promise gives
others legitimate expectations that the promisor takes her political duties seriously.
These expectations, the argument goes, demonstrate the validity of the promise.

Wolff's critics are right that the scope of what counts as a promise, including
what others can legitimate expect as a result of the promise's being made, is socially
determined. To return to an example I described in my chapter on consent, consider
what happens when I go to a bar with a friend who tells me that "If anything goes
down in the bar, I've got your back." My friend may mean something very different
than when Joe says the same thing to Steve. My friend and I implicitly take for granted
that the promise is void if I get drunk and begin to pick fights with the other patrons.
My friend meant that he would help me (to a certain extent) if someone became
aggressive toward me or if I wanted to impress someone romantically.

This might not at all be what Joe meant when he made his promise to Steve. In
fact, he may have meant precisely that he would be happy to help if Steve became
belligerent. Our respective histories – both individual and shared – establish the
implicit boundaries of the promise. Similarly, every competent member of a
community understands the social significance of some actions. These community
members may excuse an outsider for making a mistake (say, not realizing that her
actions created legitimate expectations on the part of the community members), but to
the extent that they do so, they will view this outsider as being socially incompetent.
Wolff's agent, the critic might conclude, can only reject the social significance of the promise to obey the law at the cost of announcing their own social incompetence.

This critique fails because it puts too much weight on the proposition that people understand what I say when I make a general promise to obey the law. We should read Wolff as saying that a promise to obey the state is null from the start because the vast majority of people are making a conceptual – not necessarily a moral – mistake when they consider what such a promise means.\(^{631}\) People are not speaking gibberish when they make such a promise. Most of us have a feeling that we know what someone means when they make such a promise. Yet if we were to ask these people if a promise to * (as described by Reiman) would be valid, I imagine that most would say no. So what is going on here?

Wolff's statement about our emotive attachment to the laws of our country provide a clue to resolving this tension:

> When I take a vacation in Great Britain, I obey its laws, both because of prudential self-interest and because of the obvious moral considerations concerning the value of order, the general good consequences of preserving a system of property, and so forth. On my return to the United States, I have a sense of reentering my country, and if I think about the matter at all, I imagine myself to stand in a different and more intimate relation to American laws. They have been promulgated by my government, and I therefore have a special obligation to obey them. But the anarchist tells me that my feeling is purely sentimental and has no objective moral basis.\(^{632}\)

I believe that Wolff would say that a promise to obey the law is actually a

---

\(^{631}\) It is possible to make both types of errors simultaneously. I may be led to a conceptual error by evil motives. That said, Wolff does not give us enough to speculate what proportion of people who believe in state legitimacy are being immoral – abdicating the moral responsibilities – and how many are being intellectually sloppy.

demonstrated commitment to obey the law on some other ground, usually pre-reflective emotion. Most people likely believe that we have an obligation to obey the law regardless of the promise, even if the promise strengthens the associative ties of the members of the community. But – and this is the crucial point – this means that those who say that Wolff cannot make promises have missed their mark. Wolff does not object to all promising. He objects to Reiman's 'wild-card' promises, and he (I believe) would say that promises to obey the state are a different kind of performative action than are promises to meet a friend for dinner. Political promises amount to a demonstration of one's commitment to an already existing duty.

Wolff does not believe that this duty exists. Many others do. If it does exist, however, it has nothing to do with promising and thus a different critique of Wolff altogether. If I say that the state is legitimate and I promise to obey it, a Wolffian would say that the promise is void in any case but that I am compounding my error by renouncing my status as an autonomous being by virtue of announcing (not successfully promising) my subservient status.

IV. Conclusion

IV.1 Wolff dodges a bullet

In this chapter, I sought to address and dispel the argument that Wolff's radical defense of autonomy precludes him from endorsing promising as an institution. His
conception of the person is sufficiently robust to allow for convention and cooperation. What's more, although his conception of morality is thinner than Kant's, he nonetheless would not allow people to break their promises willy-nilly. Finally, he picks up on the very plausible intuition that a promise to obey the state is not 'just another promise.' There are promises and there are promises. As we have seen, some promises simply do not "come off," and others aren't really promises as we normally understand them. For Wolff to say that promises to the state never succeed does not commit him to the position that all promises must fail.

If my arguments are sound, supporters of Wolff and a priori philosophical anarchism more generally will find their burden lessened. Those supporters still have work to do; the charges against Wolff are varied and scattered. My hope is that this chapter leads to the dismissal of one of the most prominent and disturbing of these charges. There may be hope yet for philosophical anarchists who feel that the new breed of philosophical anarchism is too heavy on the philosophy and too light on the anarchism.

IV.2 In lieu of a happy ending

In Franz Kafka’s final novel, The Castle, the protagonist “K.” spends hundreds of pages trying in vain to reach a castle which serves as the administrative center for a small town. K.’s efforts have been described as, among other things, an allegory for Christian salvation, bureaucracy, anti-Semitism, and the search for companionship.\footnote{Burrows, “Winter Read”; Spafford, “Franz Kafka’s Quest.”}
What is notable about the *Castle* is K.’s steadfastness; he moves forward but his goal remains elusive. And while I ended my previous paragraph on a suitably hopeful note, I will now end my dissertation on an ambivalent one. The castle has not been reached, even if many steps were taken. Along the way, three meta-questions became more and more difficult to ignore.

First, recall that I wrote and included this chapter as a kind of salve for those who might be frustrated with my long-winded conclusions regarding a posteriori philosophical anarchism. The first question, then, is: what can we hope for if we engage in the debate between archists and a posteriori philosophical anarchists? Is the best we can hope for the kind of ultimately unsatisfying back-and-forth that appears unlikely to sway specialists, let alone other political theorists? It’s hard to avoid getting bogged down when engaging in the debate, where theories are rarely destroyed but are often sniped at. If getting into this theoretical quagmire worth the effort?

Second, and on a similar note, what might really be gained from a turn to a priori philosophical anarchism? In this chapter I argued that Wolff dodged a bullet, but, as I said back in Chapter 1, there have been *many* bullets fired at Wolff. It is hard to blame commentators for just assuming that Wolff’s position, unintuitive and attacked as it is, is not worth seriously considering. It is hard to say whether taking up Wolff’s cause is an exercise in intellectual bravery or the worst example of academics wasting their time debating abstract points that have no hope of being relevant.

Taken together, these questions lead to the third question: who cares? This question, faint when I started my work, became increasingly hard to ignore as I
progressed and discussed my work with others. I do not mean that the fundamental questions are unimportant – they are. And those who deride philosophical anarchism for (by its own admission) having no direct political implications miss the point. Philosophical anarchists are, at their core, worried moral contamination. Everyone agrees that we should ultimately do what we think is the right thing, and the philosophical anarchist’s contribution is to try to cleanse our conscience of what they believe are common but invalid hierarchical relationships. Interpreted as charitably as possible, philosophical anarchism gives us tools to help stop us from fooling ourselves. Using this interpretation, we can say that without the efforts of philosophical anarchists and those like them, our attempts to do the right thing may result in our inadvertently perpetuating harms.

Yet the question persists, and it does so because modern philosophical anarchists appear to want to delete a core political concept without refashioning, even tentatively, a new conceptual world. Their insistence on the position that their “anarchism” does not entail a duty to oppose the state – a position that is necessary for them to avoid calling themselves standard (or “political”, as per Reiman) anarchists – prevents them from filling the void that is left by the removal of the political authority that is ever-present in the mental models of most citizens. To say “I am a modern philosophical anarchist” means something because it can be followed up with a fuller description of who one is, but “I am a modern philosophical anarchist, full stop” is hard to understand. With that in mind, I want to suggest two ways for philosophical anarchism to be an area of study that could appeal to more than those who have a
fetish for such questions. Unfortunately, each of these proposals mutilates the field to some degree, as each requires us to rethink the meaning and relevance of philosophical anarchism.

First, the study of philosophical anarchism can be used as a toolbox, not by philosophical anarchists, but by their archist opponents. If so-called philosophical anarchists want to engage in a critical enterprise, so be it. Wolff makes a strong argument against the compatibility of (a sort of) Kantian autonomy and legitimate authority? So much the better for utilitarians who oppose these kinds of autonomy-based justifications. An up-and-coming gratitude theorist could draw on existing critiques of fair play theory to justify the superiority of her pet theory. In a world where archists themselves will disagree about the true bases of legitimacy, the philosophical anarchist project can sharpen the debate by providing ammunition for all sides. They can create weapons of war.

But here saying “I am a philosophical anarchist” simply means that one engages in a critical practice. Others would be justified in wanting to know, at least in principle, which archist or anarchist subvariant she prefers and provisionally endorses. After the weapons are deployed, which perspective fares the best? They do not need to answer this question as professional philosophers who specialize in deconstruction, but the acknowledgment of the incompleteness of their project should entail a greater modesty about the meaning of labeling oneself a philosophical anarchist.

Second, philosophical anarchism can be used as a waypoint toward a political theory that provides a fuller description of our political lives, one that avoids the
vacuum left by those philosophical anarchists who effectively tell us to carry on as we see fit, minus a massive piece of our conceptual apparatus. Thoreau goes a way towards this, and this precisely why I do not consider him a standard philosophical anarchist. Putting aside his contention that we should resist clearly unjust states, a belief common among both philosophical anarchists and archists, who both start with the premise of a basically just state, Thoreau supplements his ambivalence towards a limited government with a fuller account of how one ought to live. \(^{635}\) And to be fair, both Simmons and Wolff have made some headway expanding their respective neo-Lockean and neo-Kantian positions into a fuller picture of how we ought to conceptualize and live our political lives. Yet in doing so, it is hard to avoid Reiman’s conclusion that philosophical anarchists must ultimately choose a side. Should we oppose the state, ignore it (which is a type of opposition), or act in ways that ultimately support it? Philosophical anarchists may, like Thoreau, continue to argue that the relationship should be reconceptualized, but this reconceptualization only makes sense in a broader framework that fleshes out the consequences of this move on our political and social lives.

There are of course other options available – modern philosophical anarchists could go what I jokingly call “full Stirner” and adopt a radical skepticism that questions the validity of all fixed concepts, just as Max Stirner did in \textit{The Ego and His Own}. Possibilities abound, and that is precisely the problem. What is philosophical

\(^{635}\) Recall Thoreau’s stance on voting: “I cast my vote, perchance, as I think right; but I am not vitally concerned that that right should prevail. I am willing to leave it to the majority. Its obligation, therefore, never exceeds that of expediency. Even voting for the right is doing nothing for it. It is only expressing to men feebly your desire that it should prevail.” \textit{Civil Disobedience}, para 12.
anarchism good for? If philosophical anarchism were vindicated, would anyone notice?

Anyone reading this far know that I do not believe that the problem of justifying (or denying) the legitimacy of the state can be answered simply and unequivocally. I believe that considering deliberation takes us a step further in articulating a necessarily complicated story, but even this is pointless unless it is ultimately attached to either a pragmatic goal, such as supplying others with parts for their theories, or a more far-ranging one that provides a fuller description of what the worldview that would endorse this position would look like. Otherwise, those of us who are interested in such matters risk becoming like K., forever walking with an eye toward an unreachable castle.
Bibliography


Chapman, Gary. *The Five Love Languages: The Secret to Love that Lasts*. Chicago:


Gutmann, Amy, and Dennis Thompson. *Why Deliberative Democracy?* Princeton:


Kasson, John F. *Rudeness and Civility: Manners in Nineteenth-Century Urban*


http://xroads.virginia.edu/~hyper2/thoreau/civil.html

Trump, Donald Jr. March 6, 2017.  


