RETHINKING MODERN CITIZENSHIP:
TOWARDS A POLITICS OF INTEGRITY AND VIRTUE

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by

David Thunder, B.A., M.A.

_____________________________________
Michael Zuckert, Director

Graduate Program in Political Science
Notre Dame, Indiana
April 2006
RETHINKING MODERN CITIZENSHIP:
TOWARDS A POLITICS OF INTEGRITY AND VIRTUE

Abstract

by

David Thunder

There is a malaise at large in our modern liberal democracies arising from the failure of our public life to be grounded in, and in turn inspire, our deepest moral questions and our ceaseless quest for ethical integrity or wholeness. Many current ideals of citizenship expressly or implicitly require citizens to subordinate their most cherished ethical commitments to the norms of liberal citizenship. Critics have pointed out that this amounts to a deeply problematic bifurcation of the moral life. However, they have failed to develop an alternative vision of citizenship that both accommodates integrity and responds to the challenges of a modern pluralistic polity, such as toleration and political stability. This is the principal task I set myself here, drawing on the strengths of Aristotle’s account of the virtues but also affirming the values of a broadly liberal regime.

Chapter 1 is a broad introduction to and overview of the project. Chapters 2 and 3 show that John Rawls’s “political liberal” ideal of citizenship fails both in terms of its own criteria of legitimacy and in terms of the need to accommodate the integrity of citizens. In
Chapter 4 I develop a positive account of the virtue of integrity and address some typical liberal objections against this virtue. Chapter 5 offers a deeper and broader critique of political liberalism by confronting it with the problems of respect, stability, and integrity and tracing the roots of its inadequacy to twentieth century contractualist moral theory. Finally, in chapter 6 I sketch part of an aretaic or virtue-ethical account of citizenship that merges insights from virtue ethics with insights from the liberal tradition. The aretaic account claims to be both more accommodating of integrity than contractualist models, and better equipped to understand and address traditional problems confronting a liberal regime, such as respect and political stability.
For my parents
# CONTENTS

## ACKNOWLEDGMENTS

---

## CHAPTER 1: INTRODUCTION

1.1 The Challenge of Integrity

1.2 Competing Strategies for Meeting the Challenge
   1.2.1 The Method of Avoidance
   1.2.2 The Method of Engagement

1.3 The Challenge of Citizenship

1.4 Citizenship and Integrity

1.5 Overview of the Argument

---

## CHAPTER 2: THE RAWLSIAN IDEAL OF CITIZENSHIP

2.1 Ideals of Citizenship

2.2 The Theoretical Context
   2.2.1 The Fact of Reasonable Pluralism
   2.2.2 The Concept of the Citizen as Free and Equal
   2.2.3 The Concept of Society as a Fair and Stable System of Social Cooperation
   2.2.4 The Distinction Between a “Political Conception” and a “Comprehensive Doctrine”
   2.2.5 The Distinction Between Public and Nonpublic Reason
   2.2.6 The Idea of a Well-Ordered Society

2.3 The Rawlsian Ideal of Citizenship
   2.3.1 The Place of the Ideal of Citizenship in Political Liberalism
   2.3.2 Reasonableness and the Cooperative Virtues of Citizenship
CHAPTER 3: A CRITIQUE OF PUBLIC REASON FROM INTEGRITY AND RECIPROCITY ................................................75

3.1 The Duty of Civility .................................................................78
3.1.1 Content of the Duty: Public Reason ................................78
3.1.2 Grounds of the Duty: Reciprocity ..................................83

3.2 A Rawlsian Argument Against the Duty of Civility .................92
3.2.1 Should We Subject the Duty to the Test of Reciprocity? ....92
3.2.2 Who Are the “Reasonable” Citizens to Whom the Duty Must Be Justified? .......................................................94
3.2.3 Does the Duty of Civility Pass the Test of Reciprocity? ......101

3.3 The Critique from Integrity .....................................................110
3.3.1 Integrationist and Non-Integrationist Citizens ...............112
3.3.2 Reasonable vs. Unreasonable Integrationist Citizens .......115
3.3.3 The Likely Impact of the Duty on Reasonable Integrationist Citizens .................................................................117

3.4 Implications of the Duty for Political Stability .......................132

3.5 Rethinking Reciprocity .........................................................135

CHAPTER 4: THE CONCEPT OF INTEGRITY ..................................139

4.1 A Preliminary Exploration of Integrity ................................143

4.2 The Subject of Integrity ..........................................................146

4.3 Existing Accounts of Integrity .................................................150

4.4 A Critique of Existing Accounts ..............................................156

4.5 Integrity as a Family of Virtues ..............................................161
4.5.1 The Relationship Between Integrity and Virtue ..............161
4.5.2 The Virtues of Integrity .....................................................169
4.5.3 Connecting the Form and Substance of Integrity ..........177
4.5.4 Integrity as a Matter of Degrees .......................................179

4.6 Objections and Replies ..........................................................180
# CHAPTER 5: THE LIMITS OF CONTRACTUALIST ACCOUNTS OF CITIZENSHIP

5.1 Some Criteria for a Successful Account of Citizenship .................. 208

5.2 Why We Need a Radical Alternative to the Rawlsian Ideal of Citizenship .......................................................... 212

5.2.1 Respect for Persons ............................................................... 213

5.2.2 The Problem of Stability ....................................................... 218

5.2.3 The Problem of Integrity ....................................................... 221

5.3 The Limits of Contemporary Contractualism ......................... 222

5.3.1 John Rawls: Political Liberalism ............................................. 222

5.3.2 Iris Young: Deliberative Democracy ...................................... 226

5.3.3 Chantal Mouffe: Agonistic Pluralism ..................................... 232

# CHAPTER 6: TOWARDS A POLITICS OF INTEGRITY AND VIRTUE: THE ARETAIC IDEAL OF CITIZENSHIP

6.1 Personhood, Citizenship, and the Good Life ......................... 239

6.1.1 The Concept of Person and Personal Virtue .......................... 244

6.1.2 Persons and Social Roles ..................................................... 248

6.1.3 Personal Virtue and the Role of Citizen ................................. 256

6.2 The Aretaic Ideal of Citizenship Applied to Civic Discourse ...... 263

6.2.1 The Conditions Marking Political Deliberation in Liberal Democratic Societies .................................................. 265

6.2.2 Goals and Virtues of Good Deliberation ................................. 269

6.2.3 Deliberative Civic Virtue and Integrity ..................................... 282

6.3 Comparing the Merits of Aretaic and Contractualist Ideals of Citizenship .......................................................... 284

6.3.1 Respect for Persons .............................................................. 286

6.3.2 Political Stability ................................................................. 288

6.3.3 The Demands of Integrity ..................................................... 292

6.4 Objections and Replies .......................................................... 296

# CONCLUSION ............................................................................. 309

# BIBLIOGRAPHY .......................................................................... 314
One of the most valuable lessons this project has taught me is that in undertaking any serious intellectual inquiry, one incurs more intellectual and moral debts than one can hope to ever acknowledge, let alone repay. If Aristotle’s magnanimous man loathes being in the debt of his peers, then perhaps Aristotelian magnamity is a virtue best displaced by a humble acknowledgment of one’s extreme dependency on others.

I will mention some of the debts I have accumulated during this project, but certainly not all. First and foremost, to God, for bringing me into existence and blessing me with the talents, opportunities, friends, and mentors necessary to bring this project to completion. Second, to my parents, for bringing me into this world and teaching me more about virtue and integrity than any book could teach me. Third, to my family and friends, near and far, for their love, good humor, and support throughout my time as a doctoral student. And fourth, to my colleagues and mentors, especially but not exclusively at Notre Dame, for their encouragement and assistance, whether in my general education or in helping me think through the specific set of issues centered around citizenship and integrity. Thanks are due in particular to several of my friends with whom I have had extensive conversations on aspects of the project, especially James Helmer, Matt
Mendham, Michael Siebert, Miguel Rumayor, Paolo Monti, Christian Miller, Chris Eberle, Carmen Pavel, Micah Schwartzman, and Fevzi Bilgin. Fifth, to the University of Notre Dame and the Earhart Foundation, for providing financial support to enable me to pursue this research.

I would like to extend a very special note of thanks to my advisor, Michael Zuckert, who has gone beyond the call of duty in supporting my intellectual and professional development and in playing a challenging yet entirely constructive role in pushing my ideas along. All of my committee, Michael Zuckert, Mary Keys, Fred Dallmayr, and David Solomon, have been very generous and supportive throughout this process.

I thank all my teachers and mentors, but I would like to mention a few by name: The late Phil Quinn, who lived long enough to read and comment on a draft of my dissertation proposal. The small seminar I took with him on religion and politics was an apprenticeship in careful and exacting yet rich and charitable thinking. Alasdair MacIntyre, who graciously agreed to read and discuss the introduction and chapter 4 (“The Concept of Integrity”). Paul Weithman, who kindly gave me detailed comments on my critique of Rawls in chapter 3. Nicholas Wolterstorff, who discussed aspects of the project with me at some length, and agreed to read and comment on chapter 4. Last but not least, I received excellent feedback from anonymous reviewers at the *American Journal of Political Science* on part of chapter 3. The internal argument against the duty of civility essentially reproduces an article forthcoming at the *AJPS*, “A Rawlsian Argument Against the Duty of Civility” (July 2006).
I have presented versions of parts of the dissertation at various colloquia and conferences, which provided me with invaluable feedback to refine the project along the way. I presented versions of chapters 3, 4, 5, and 6 to the Notre Dame graduate ethics colloquium and I would like to thank the participants for their thoughtful contributions, especially Peter Wicks, David DiQuattro, Raymond Hain, Becky Stangl, Sean Walsh, Holly Vandewell, Paolo Monti, Eric Gregory, Robert Audi, David Solomon, Lynn Joy, and Kevin McDonnell. I presented versions of chapter 3 to the Association for Political Theory at Calvin College (10/17/2003), and the Political Science Graduate conference at Marquette University (3/19/2005), and a version of chapter 4 to Notre Dame’s Political Theory brownbag series (10/08/2004). Finally, I presented a compressed version of chapters 5 and 6 to the Association for Political Theory at Washington University, St. Louis (10/22/2005).

There are many intellectual debts that will remain unacknowledged: who knows what book I have read, or what conversation I have had, or what suggestions I have been offered, that have set me off in this or that direction? But a few that I feel compelled to acknowledge are Aristotle, whose common sensical approach, though seemingly tedious at times, is full of surprising insights; Thomas Aquinas, who, though barely mentioned in this work, continues to quietly form my intellectual habits; Alasdair MacIntyre, whose work taught me what it means to think within a tradition and across traditions; Charles Taylor, who taught me how one can recognize the limitations in modernity yet approach it with a tempered optimism; Elizabeth Anscombe and Bernard Williams, who taught me
that ethics is primarily about how to live, not how to be “moral”; Rosalind Hursthouse, whose refreshingly practical defense of virtue ethics has left a deep mark on my own grasp of the virtues; and Danielle Allen, who helped me imagine how an aretaic approach to liberal democracy might speak eloquently to practical political problems.

University of Notre Dame, 9 April 2006
CHAPTER 1: INTRODUCTION

In recent decades, political philosophers have shifted their attention from the legal bases of citizenship to its social and moral foundations. Something like an ideal of citizenship, i.e. a desirable pattern of behavior and dispositions,\(^1\) has been advanced, sometimes inchoately, other times more systematically, by a host of liberal and democratic theorists from John Rawls to Jürgen Habermas.\(^2\)

Underpinning this turn towards ethical ideals of citizenship is the realization that the political and legal order is parasitic upon a set of dispositions or virtues that are not themselves guaranteed by the design of political and legal institutions.\(^3\) A thriving liberal

\(^1\) Talk of “ideals” of citizenship may suggest what philosophers often speak of as “superogatory” or highly aspirational ideals, whose fulfillment may be viewed as going beyond the call of duty, or even as heroic. But for our purposes, an “ideal” will simply mean a desirable pattern of action and dispositions, which thus far implies neither strict moral duty nor superogatory virtue.


\(^3\) For a thorough account of these dispositions as interpreted by Hobbes, Locke, Kant, and Mill, see Peter Berkowitz, *Virtue and the Making of Modern Liberalism* (Princeton, New Jersey: Princeton University Press, 1999).
democratic order seems to require some publicly recognizable ideal or ideals of civic excellence besides the familiar rights and duties of citizenship established by law.

§1.1 The Challenge of Integrity

Yet the project of articulating a viable ethic of citizenship that is supportive of a broadly liberal political order, important as it is, seems problematic at first blush: if we accept that modern Western polities are characterized by a plurality of reasonable comprehensive doctrines or conceptions of the good, then to hold citizens to ideals of civic excellence extending far beyond the strict requirements of law, to expect them to regulate their behavior and attitudes with reference to a single publicly acknowledged ideal of citizenship, seems on its face not only unrealistic, but potentially violative of their freedom to pursue a conception of the good of their choice within the bounds of law and public order.

Nonetheless, something like an ideal of citizenship is already operative in liberal democracies such as the United States, Canada, Britain, France, and Germany. There is a set of expectations, norms, and excellences associated with citizenship, however fluid or open to divergent interpretations it may be. And there clearly are exemplars of good citizenship in the public political cultures of such countries.

Furthermore, these publicly acknowledged ideals of citizenship are shared by people of very different religious and moral perspectives. Therefore, rather than rejecting the very notion of a modern ethical ideal of citizenship as somehow oxymoronic, any effort to articulate such an ideal in modern circumstances should proceed fully cognizant of the delicate nature of the task.

---

4 In the United States, for example, exemplars of civic excellence include Abraham Lincoln, Martin Luther King, Jr., and more recently, the firemen who came to the aid of victims of the September 11 attacks on the World Trade Center in New York.
The fundamental challenge for a public ideal of citizenship is to be sufficiently robust so that it is politically consequential and supportive of a healthy constitutional regime, while sufficiently adaptable so that it can be incorporated into diverse ways of life without doing violence to people’s deepest moral commitments. Anyone who hopes to advance a viable ideal of citizenship today must provide some grounds for the hope that it is indeed possible to be at once a good person and a good citizen, that the two do not necessarily come apart. Let us call this need to reconcile good citizenship with human goodness simpliciter—and we need not assume that the need is universally felt—the challenge of integrity. A rough preliminary definition of integrity will suffice for now: the ideal of integrity demands a conscientious effort to order one’s commitments, desires, evaluations, and dispositions in accord with what one responsibly takes to be the good. Such an effort aspires towards the mutual coherence of a person’s diverse ethical and moral commitments.

§1.2 Competing Strategies for Meeting the Challenge

§1.2.1 The Method of Avoidance

Now, there are two principal strategies one might adopt in attempting to meet the challenge of integrity. The first strategy I will call the “method of avoidance.” The idea

---


6 I appropriate the term from Rawls (see John Rawls, "Justice as Fairness: Political Not Metaphysical," Philosophy and Public Affairs 13, no. 3 (1985), 230-231). Although Rawls does not frame the method of avoidance explicitly as a strategy for meeting the challenge of integrity but as a way of publicly justifying a political order marked by a plurality of reasonable comprehensive doctrines, I believe Rawls’s method of avoidance can be plausibly interpreted as a strategy for responding to the problem of integrity.
is to advance an ideal of citizenship that is as independent of controversial or idiosyncratic conceptions of the good as possible, and that forms into a coherent whole moral ideas or intuitions implicit within the public culture.\(^7\) This ideal is then offered to citizens in a “free-standing” way, that is, independent of any deep account of morality or the good, in the hope that they can integrate it as a “module” into their existing system of beliefs.\(^8\) Using this strategy, a variety of “conjectures” or hypotheses can be advanced about how the ideal might be sanctioned or permitted by the rest of the person’s moral beliefs or values, but the details are left up to individuals to work out in light of their individual “comprehensive views.” The hope is that an ideal that self-consciously eschews deep moral or metaphysical premises, and is inspired by broadly liberal principles implicit in the public culture, should be acceptable to most citizens upon due reflection. No elaborate account of integrity or the process of moral integration is necessary. On the contrary, any such account would almost certainly burden the ideal of citizenship with controversial conceptions of the moral life, and as such, should be avoided.\(^9\)

\(^7\) In speaking of a “public culture,” we need not assume any bright line between the public and nonpublic culture, only that there are ideas that clearly fall within one or the other.

\(^8\) I am borrowing the notion of a “module” from John Rawls, *Political Liberalism, The John Dewey Essays in Philosophy* (New York: Columbia University Press, 1996), 12-13. From now on, this work will be abbreviated to PL.

\(^9\) This is the approach advocated by John Rawls. See, for example, PL, 10-11: “Political liberalism, then, aims for a political conception of justice as a freestanding view. It offers no specific metaphysical or epistemological doctrine beyond what is implied by the political conception itself…citizens themselves, within the exercise of their liberty of thought and conscience, and looking to their comprehensive doctrines, view the political conception as derived from, or congruent with, or at least not in conflict with, their other values”; and PL, 12: “I assume all citizens to affirm a comprehensive doctrine to which the political conception they accept is in some way related.”
§1.2.2 The Method of Engagement

A second strategy for meeting the challenge of integrity rejects the “method of avoidance” of the first strategy, contending that no political theory with any normative bite can refrain from engaging weighty philosophical and moral questions, if not overtly then in its implicit commitments. But an equally important and sometimes overlooked point is that a “free-standing” ideal of citizenship, while it may contingently turn out to be acceptable to many citizens, cannot speak convincingly to those citizens who have serious doubts about how the role of citizen and its attendant values and purposes relate to their other roles and values, and aspire to some sort of overall coherence among their political, moral, and/or religious commitments. In other words, even assuming that a free-standing ideal of citizenship speaks to the person qua political actor\textsuperscript{10}, it fails to speak to the person qua person. And this failure is fatal to its chances of winning the free allegiance of citizens who are concerned primarily with living good moral lives and only secondarily with being good citizens.\textsuperscript{11}

Having rejected the “method of avoidance” of the first strategy, the second strategy for meeting the challenge of integrity aims to speak to the whole person, and not just to the citizen. On this account, citizenship is one important role among many that can

\textsuperscript{10} Although it is not clear what it would mean for an explanation to speak to me as a political actor and not as a person. If, for example, someone were to say to me, “As a citizen, you are obligated to submit to political authority, even though it is ordering acts that are evil in your eyes as a person, since your personal point of view does not count in your role of citizen,” this appears to attribute a degree of autonomy to the role of citizen that entirely eclipses full personal agency. But this will be discussed in more detail in chapter 4, “Integrity.”

\textsuperscript{11} In John Tomasi, \textit{Liberalism Beyond Justice: Citizens, Society, and the Boundaries of Political Theory} (Princeton and Oxford: Princeton University Press, 2001), Tomasi argues persuasively that the case for political liberalism is less than compelling until political liberals have addressed the consequences of their theory for the personal lives of citizens. Though Tomasi is right on this point, I believe his attempt to reduce the conflict between political liberalism and the personal lives of citizens while preserving the “architecture” of public reason essentially intact, does not go far enough to accommodate personal integrity. Cf. my more radical critique of public reason from integrity in §3.2 and 3.3 below.
only be adequately explained and justified in the context of an interpretation of the moral life more broadly construed. The challenge of integrity can only be adequately met by developing a fuller theory of integrity and showing how some determinate ideal of citizenship can be viewed as congruous with its demands.

Call this second strategy the “method of engagement.” This is the strategy I favor, for reasons that should become apparent in the course of its elaboration. It is worth commenting briefly on what is at stake in our choice between these two rival strategies. Underlying the method of avoidance is the belief that moral theory and political philosophy are two autonomous disciplines, with their own distinctive methods and goals. This is not to suggest that political philosophy is completely separate from morality, but it is to suggest that the scope of political theory and the goals that inform it are much more modest and constrained than those of moral theory. Whereas moral theory seeks to understand the grounds and content of morality in general, political theory seeks to understand the basis of political order, which is a relatively constrained endeavor and does not depend on any general account of moral truth or human nature.

Along similar lines, Dworkin favors what he calls the “strategy of continuity,” which argues that liberals should “connect ethics and politics by constructing a view about the nature or character of the good life that makes liberal political morality seem continuous rather than discontinuous with appealing philosophical views about the good life.” See Ronald Dworkin, “Foundations of Liberal Equality,” in *The Tanner Lectures on Human Values* (Salt Lake City: University of Utah Press, 1990) (6). This similarity notwithstanding, Dworkin reaches very different conclusions about the role of citizen and its demands than I do. A more recent work arguing against the rigid constraints of scope Rawls places upon political liberalism is Tomasi, *Liberalism Beyond Justice*, esp. chap. 1. A good summary of the relevant thesis can be found in the introduction at xvi: “I aim to demonstrate that the normative domain of liberal theory construction is importantly wider than the domain of public, deliberative value. For political liberals in particular, there is important work for liberal theorists to attend to beyond their current fixation on questions of legitimacy and justification.” Thus far, I agree with Tomasi, but believe he cannot adequately address the public-private “interface” questions he raises without radically transforming the theoretical framework of political liberalism, in particular what he calls the “architecture of public reason.”

This separation between political theory and moral philosophy is one of the foundational commitments of political liberalism and is most clearly articulated by Rawls. See, for example, Rawls, "Justice as Fairness: Political Not Metaphysical."
The method of engagement, on the other hand, while acknowledging the distinct scopes and goals of moral and political theory, sees the relation between them as much more intimate and mutually implicating than the method of avoidance does. First, any interesting and informative political theory, whether it wants to or not, necessarily presupposes some moral theory, however inarticulate. Rawlsian liberalism, for example, is inspired by a contractarian or contractualist view of morality. Secondly, political theory is driven by its own exigencies to take moral positions often debated within moral theory. For example, while political theory proposes principles for the governance of political life and seeks principally to understand man as a political animal rather than simply as man, the political theorist must concern herself with the possible effects of political order, for good or ill, on the well-being of persons viewed as persons. She must come to some views about the human good, at least substantial enough to guide political action. Moral inquiry in a political context is not conducted in quite the same way as it is by the moral theorist as moral theorist, since it is conducted in the service of political

---

14 I argue this point at greater length in §5.3 below. For now, it should suffice as an example of how the “engaged” theorist views those who claim to adhere to the method of avoidance. Rawls extends his gratitude to T.M. Scanlon, one of the architects of contemporary contractarian morality, in the introduction to PL “for numerous instructive discussions from the beginning about political constructivism” and “for discussions about the distinction between the reasonable and the rational, and how to specify the reasonable for the purposes of a political conception of justice” (xxxii). Rawls acknowledges that his description of reasonable persons and Scanlon’s principle of moral motivation are “closely connected” (PL, 49, fn. 2). Cf. T.M. Scanlon, “Contractualism and Utilitarianism,” in Utilitarianism and Beyond, ed. Bernard Williams (Cambridge: Cambridge University Press, 1982); for a more comprehensive statement, see T.M. Scanlon, What We Owe to Each Other (Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press, 1998). Presumably Rawls would argue that the contractarian basis of political liberalism is something any reasonable person could accept. But I view contractarianism as just one more contestable theory of morality, which owes its appeal in part to its strong roots in the liberal tradition and its general affinity with modern moral sensibilities.

15 Of course, political liberals such as Rawls believe that we can confine our view of the human good to a “thin” theory of the good that can be embraced by all or most reasonable citizens. However, the claim that a thin theory of the good adequately articulates the political basis for human well-being is itself a disputed point in moral theory.
theory; however, the notion that moral and political theory are autonomous in any strong sense is not tenable.\textsuperscript{16}

The problem of integrity—that is, the problem of reconciling one’s civic commitments with the pursuit of a good human life—is a problem that arises out of any serious reflection upon political order in modern industrialized societies. It can thus be thought of as a problem internal to political theory. Yet it cannot be addressed in any politically adequate way, without recourse to some moral theory of integrity. Whether we choose to describe this as an incursion into moral theory, or instead as a part of political theory proper, is of little consequence: all that matters for my purposes is that we accept the legitimacy of the endeavor and do not judge it extraneous to the concerns of political theory.

§1.3 The Challenge of Citizenship

So far, I have offered a rough preliminary statement of the challenge of integrity confronting contemporary discussions of citizenship and indicated my approval of the “method of engagement” over the “method of avoidance” as a general philosophical approach to the problem. It is now time to say something about the challenge of citizenship and political order that confronts any strong ideal of integrity.\textsuperscript{17} This challenge has yet to be met, let alone fully acknowledged, by political proponents of

\textsuperscript{16}Thus, I agree wholeheartedly with William Galston that the project of a “free-standing” political theory, or a political theory that eschews deep ethical commitments is a non-starter (William Galston, \textit{Liberal Pluralism: The Implications of Value Pluralism for Political Theory and Practice} (Cambridge: Cambridge University Press, 2002), 8, 39-47).

\textsuperscript{17}By a “strong” ideal, I mean one that does not generally subordinate integrity to the demands of a political conception of justice or to the demands of the political order.
integrity. Indeed, such accounts tend to be more reactive than responsive to the liberal ideal of politics and citizenship. That is to say, they successfully expose problematic aspects of the liberal ideal but do not go very far in positively addressing the concerns that have led liberals to rigidly separate the public from the private spheres of morality.

Political proponents of integrity have argued that in making the demands of the political domain not only weighty but virtually overriding, and in demanding that citizens bracket their most fundamental convictions during some or all processes of civic deliberation, the liberal ideal of citizenship rules out the pursuit, not to mention the possible attainment, of integrity or coherence between the diverse commitments that inform a person’s life. Now, the critique from integrity has exposed an important vulnerability in the liberal account of citizenship: its failure to adequately acknowledge, let alone accommodate, the problem of integrity. However, proponents of integrity have tended to concentrate their energies on refuting the liberal model of citizenship rather than developing a positive alternative of their own. Furthermore, they have been less than forthcoming in spelling out what the notion of integrity amounts to, how it can be responsive to the legitimate demands of a political community, and what role, if any, the

---

18 By political proponents of integrity, I mean those who advocate the accommodation of integrity by political authority and self-consciously deny the Rawlsian-liberal attempt to cordon off comprehensive doctrines from political conceptions of justice. As we shall see shortly, there is also a body of literature within moral philosophy on the concept of integrity, but this literature does not treat in any sustained way the relation between integrity and citizenship or political order. One exception to this neglect is Mark Philp, “Citizenship and Integrity,” in Integrity in the Public and Private Domains, ed. Alan Montefiore and David Vines (London and New York: Routledge, 1999).

19 One of the clearest statements of this critique, albeit one that focuses exclusively on the case of religious citizens, can be found in Wolterstorff’s critique of Rawlsian liberalism in Robert Audi and Nicholas Wolterstorff, Religion in the Public Square: The Place of Religious Convictions in Political Debate (London: Rowman & Littlefield Publishers, Inc., 1997). Arguments along similar lines but still focusing on the case of religious citizens can be found in Eberle, Religious Convictions in Liberal Politics. Finally, arguments framing the critique from integrity in more general moral terms can be found in Michael Perry, Morality, Politics, and Law (New York and Oxford: Oxford University Press, 1988) and Dworkin, “Foundations of Liberal Equality.”
notion of integrity might play in the moral life of non-religious persons. In short, political advocates of integrity have neither sufficiently elaborated the concept of integrity to bring to light its full political implications, nor developed a realistic and attractive vision of citizenship and civic life within which integrity can find a home.

In light of this silence about the content and positive implications of integrity for political life, it seems likely that many will view the failure of the liberal ideal to accommodate integrity not as a fatal flaw but as the price we have to pay for an ideal of civic life that is both compatible with government by consent, and conducive to public order. After all, if citizens insist on forging political justifications overtly grounded on their own idiosyncratic “comprehensive doctrines” or worldviews, surely they are showing disrespect to their fellow citizens in refusing to justify coercion on grounds that are reasonably accessible to others with different worldviews? In addition, if each person calls upon whatever justificatory resources he happens to find personally persuasive when he engages in public political discourse, this seems likely to reduce political discourse to bald assertions of truth with little or no hope of public settlement. Discursive anarchy may well issue in political anarchy or at least severe public disorder. In short, those who are already committed to the principle of government by consent as well as the high priority of political stability may view integrity as an ideal that, while attractive in the abstract, is too costly to maintain in practice.

---


21 Here, by “government by consent,” I do not mean merely government by consent of a majority, but government by the consent of all citizens committed to seeking to live with others on fair terms of cooperation.
§1.4 Citizenship and Integrity

Both the challenge of integrity to liberal ideals of citizenship and the challenge of citizenship to ethical ideals of integrity seem on their face so compelling that we may be tempted to think that we must choose between integrity and citizenship. But such a verdict would probably be premature, since there has not as yet been any sustained effort to reconcile within one account a robust ideal of integrity and a morally substantive ideal of citizenship. It is only by developing the concepts of citizenship and integrity with the challenges on both sides in full view, that there is any hope of breaking the impasse and articulating an ideal of citizenship that is genuinely accommodating of the demands of integrity and an ideal of integrity that is responsive to the demands of political order.

Therefore, my argument, not surprisingly, will attempt to bring the concepts of integrity and citizenship, and their associated requirements, into a dialectical encounter. The reason for believing that such a dialectical encounter is likely to be fruitful is that these two ideals do not refer to two entirely autonomous phenomena, but express different aspects of the human good. In short, both speak in some way to the fundamental ethical question, “How am I to live my life?” In other words, an investigation of integrity and citizenship should uncover different aspects of the human good. As such, it seems likely that interpretations of these two dimensions of the human good will influence each other in a dialectical manner. By this, I mean quite simply that reflection

---

22 Galston’s articulation of political order is uniquely sensitive to the challenge of integrity, and his concept of “expressive liberty” expresses one important aspect of integrity. However, there is little sustained discussion of integrity as a virtue, and its positive relation to the virtues of citizenship. See Galston, Liberal Pluralism, esp. 110-123.

23 On the primacy of this question in the moral life relative to questions such as “what is my duty?” or “what will make for the greatest happiness of the greatest number?” see Bernard Williams, Ethics and the Limits of Philosophy (Cambridge, Massachusetts: Harvard University Press, 1985), esp. 1-21.
up upon the question of good citizenship is likely to inform our interpretation of integrity in significant ways, and vice versa. For example, an ideal of integrity should not in principle rule out all attractive ideals of citizenship any more than an ideal of citizenship should in principle rule out all attractive ideals of integrity: if our preliminary definition of integrity rules out the role of citizen even in a political regime that is optimally just, for instance, we may be driven to reconsider our definition of integrity. In this way, there is a kind of mutual “checking” of the concepts of integrity and citizenship.\textsuperscript{24}

The provisional goal is to begin to elaborate ideals of integrity and citizenship that hold out some hope of coexistence, however fragile, rather than mutual destruction. This goal of congruence will serve as a sort of “regulative ideal” for the overall project, which is not to say that it is an unshakable, dogmatic commitment, but that I will make every reasonable effort to achieve it. If the resulting account proves untenable or contrived, then it may well be that the congruence I seek is unrealistic or deeply misconceived. Nonetheless, the effort itself seems likely to yield insights into moral agency under modern conditions, and for this reason if no other, seems worthwhile whether ultimately successful or not.

\textbf{§1.5 Overview of the Argument}

The basic intuition driving this project is that contemporary liberal accounts of citizenship are insufficiently sensitive to the challenge of integrity, while political proponents of integrity, for their part, are insufficiently sensitive to the demands of

\textsuperscript{24} This resembles Rawls’s notion of “reflective equilibrium” between considered judgments. Here, the equilibrium is not between general principles and particular judgments but between the minimal requirements of integrity and citizenship respectively, assuming that both concepts express legitimate values within the well-lived life. For a nice summary of Rawls’s idea of reflective equilibrium, see PL, 8.
citizenship and political order. Curiously, what unites both sides is the absence of any systematic treatment of the concept of integrity. This work is an attempt to fill this gap in the conviction that no viable ideal of citizenship can be developed without grasping the fundamental value of integrity in the moral life. The main goal of this work is to show that integrity, properly understood, is a reasonable aspiration compatible with a broadly liberal democratic political order. Insofar as I offer an ideal of citizenship, it will be largely in the service of this goal. Therefore, rather than present a comprehensive or fully developed account of citizenship, I present a partial ideal of citizenship intended as a positive illustration of the prospects for some sort of reconciliation between the wholehearted pursuit of the good by citizens on the one hand, and their responsible participation in political life on the other.

Since I take John Rawls to be the most influential and paradigmatic exponent of the standard liberal ideal of citizenship, I will begin by laying out the Rawlsian ideal of citizenship, drawing mainly on *Political Liberalism* and subsequent works (chapter 2). The good citizen, on this view, is one who possesses the cooperative virtues of reasonableness and reciprocity, which implies, among other things, a willingness to offer one’s fellow citizens justifications of law and policy (touching constitutional essentials and matters of basic justice) that are presented in purely political terms, that is, independently from comprehensive views.

In Chapter 3, I critically evaluate the Rawlsian ideal both on its own terms and in terms of its implications for personal integrity. I argue that the Rawlsian ideal of citizenship is both internally incoherent and constitutes a formidable threat to citizens’ integrity or conscientious pursuit of the good. My internal critique of political liberalism
focuses on the so-called “duty of civility,” which requires citizens to rely upon public reason(s) while debating constitutional essentials and matters of basic justice. I argue that the principle of reciprocity at the root of the duty is ultimately self-defeating since it cannot be justified to any reasonable citizen—at least not in the way Rawls formulates it. My critique from integrity takes a number of cases of ostensibly reasonable citizens and shows that they would be unreasonably and unnecessarily impeded in their pursuit of the good by the duty of civility.

In chapter 4, I take the case for accommodating integrity further than most of its proponents by developing a positive account of integrity and defending it against some of its most serious liberal rejoinders. In this way, I aim to close a gap in the political literature on integrity by bringing integrity face to face with the challenge of politics and citizenship. On this account, a person possesses integrity in the degree to which (a) his highest-order desire is to do the good; (b) that desire is effective and stable enough to dynamically structure and re-structure his moral life (beliefs, actions, evaluations, desires, dispositions) over time; and (c) his beliefs, actions, evaluations, and dispositions are mutually coherent. In responding to concerns about the implications of integrity for political order, I hope to show that the case for the compatibility of integrity with responsible citizenship is prima facie a strong one, but the more decisive case must await my elaboration in chapter 6 of a positive alternative to Rawls’s ideal of citizenship.

---

25 Some authors who have criticized the deleterious effects of political liberalism for the ethical integrity of citizens include Wolterstorff in Audi and Wolterstorff, Religion in the Public Square; Perry, Morality, Politics, and Law; and Dworkin, “Foundations of Liberal Equality.” As we will see in chapter 4, there is a fairly extensive philosophical literature on integrity but it is virtually ignored by political theorists. I hasten to add that this is not a matter of negligence but a typical consequence of disciplinary fragmentation.
In chapter 5, I move to a deeper and broader diagnosis of the failure of political liberalism to propose a viable ideal of citizenship for modern pluralistic societies. The first two sections of this chapter (§5.1 and §5.2) aim to spell out more explicitly three central limitations of Rawls’s ideal of citizenship, namely its failure to propose (i) a viable concept of respect, (ii) a convincing account of political stability, and (iii) a sensitive and cogent treatment of the problem of integrity. The third section of the chapter broadens my critique of political liberalism by suggesting that these general limitations are not peculiar to political liberalism, but naturally occur in the conceptual world of twentieth century contractualism, of which political liberalism is just one paradigmatic instance. I compare political liberalism with two rival theories, deliberative democracy and agonistic pluralism, and find that in spite of their differences, all three theories embody ideals of citizenship whose major weaknesses can be explained by their contractualist presuppositions.

In chapter 6, building on my critique of contractualism and my positive account of integrity, I sketch part of an aretaic or virtue-ethical ideal of citizenship, focusing on the virtues required for successful and respectful political deliberation. I argue that my ideal is superior to contractualism insofar as it is both markedly more successful at accommodating citizens’ integrity, and proposes a more morally compelling and true-to-life account of the concept of respect and the conditions of political stability than Rawls or other contractualists have offered. In short, chapter 6 argues that an aretaic ideal of citizenship goes a long way towards meeting both the challenge of citizenship and the challenge of integrity within a broadly liberal political order.
Achieving congruence between the role of citizen, the other roles we occupy, and our diverse moral commitments is no easy task. Nor is demonstrating the reasonableness and political practicability of the ideal of ethical congruence or integrity. However, I hope in this introduction to have at least impressed the reader with the force and political relevance of the challenge of integrity, and the surprising fact that integrity is one of the concepts most neglected (or perhaps taken for granted) by political philosophers. I also hope to have made a good case for the inadequacy of any ideal of citizenship that refuses to, or is incapable of, speaking to the wider moral aspirations of citizens, in particular the aspiration towards integrity. In attempting to bring ideals of integrity and citizenship into a dialectical encounter, this entire project is inspired by the hope that we can answer in the affirmative the question, “is it possible to be at once a good person and a good citizen?”
CHAPTER 2: THE RAWLSIAN IDEAL OF CITIZENSHIP

In this chapter, I aim to explicate as fairly and concisely as possible the ideal of citizenship latent in John Rawls’s “realistically utopian” framework for a liberal political order. My analysis is concerned with Rawls’s mature position, meaning the position he articulated in Political Liberalism (1993) and subsequent works. As such, insofar as I draw on A Theory of Justice (1971), I do so only where I view it as fairly uncontroversially consistent with the Rawls of political liberalism. My aim in doing so is to largely prescind from debates about the relation between A Theory of Justice and Political Liberalism, lively and interesting as they are.²⁶

The general problem motivating Rawls’s work is not, “How ought a good citizen behave?” but “How is it possible to maintain a just and stable political order amidst a population marked by a plurality of conflicting and reasonable comprehensive doctrines?”²⁷ Therefore, it is not surprising that Rawls does not present any comprehensive or focused exposition of the ideal of citizenship at work in his theory. Instead, various aspects of the ideal are presented and developed at different stages of his work, sometimes roughly, other times in greater detail. I will try to gather together these


²⁷ In Rawls’s words, “how is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?” (PL, 4).
different elements into a single account of citizenship, and will allude to other aspects of Rawls’s theory only insofar as these shed light on his ideal of citizenship.

The exposition of the Rawlsian ideal will proceed in three stages: (§2.1) first, I will clarify what is meant by an ideal and dispel some potential misunderstandings of the term; (§2.2) second, I will present as concisely as possible the fundamental elements of Rawls’s vision of political order that provide the essential theoretical context for understanding his ideal of citizenship; and (§2.3) finally, I will present the ideal itself, attempting to remain as faithful as possible to what Rawls actually says rather than extrapolating generously from the text.

§2.1 Ideals of citizenship

In order to explain what I (and I think Rawls) mean by an “ideal” of citizenship, I will begin by considering some definitions of “ideal” offered by Webster’s New College Dictionary:

1. A conception of something in its absolute perfection.

2. One considered to be a standard of perfection or excellence and worthy of imitation.

3. An ultimate object of endeavor.

4. An honorable or worthy aim or principle.28

Clearly, the first definition, “a conception of something in its absolute perfection,” is too strong for what Rawls has in mind. One of the main concerns motivating Rawls’s

project is to find an adequate moral basis for political order that all can share. It would indeed be surprising if citizens who disagree deeply about ideals of human goodness in nonpolitical spheres (and Rawls would be the first to admit that this disagreement is the “inevitable outcome of free human reason” in a pluralistic society\(^{29}\)) could be brought to agree on or publicly acknowledge “a conception of citizenship in its absolute perfection”!

Now, let us consider the second definition, “One considered to be a standard of perfection or excellence and worthy of imitation.” Already, this is more promising. Rawls’s ideal of citizenship is a pattern or paradigm of excellence(s) pertaining to the role of citizen and the one who embodies it is indeed considered “worthy of imitation.” On the other hand, an ideal of citizenship, while specifying certain ways in which a citizen may excel, need not be comprehensive in the sense of specifying all ways in which a person may excel in civic life. Insofar as it is less than comprehensive, Rawls’s ideal of citizenship, while presenting a standard of excellence, by no means presents a standard of perfection.

If we accept that an ideal of citizenship specifies certain excellences associated with the role of citizen, there are still some thorny questions about the normative status or implications of such an “ideal.”\(^{30}\) For example, do we expect our fellow citizens to strive

\(^{29}\) PL, 37. Elsewhere, he observes that “many of our most important judgments are made under conditions where it is not to be expected that conscientious persons with full powers of reason, even after free discussion, will all arrive at the same conclusion” (PL, 58).

\(^{30}\) Weithman has a helpful discussion of this issue in Religion and the Obligations of Citizenship, 186-187. There is a brief discussion of the normative implications of ideals in Audi, Religious Commitment and Secular Reason, 85-86. According to Audi, we may be required by an “ideal” to refrain from doing something and yet have a moral “right” to do it. For example, although Audi maintains that citizens may be criticized for basing their votes on a religious ground (since this violates his ideal of citizenship), he does not deny that there is a “[moral] right to [do so]” (114). On my interpretation, however, it is incoherent to attribute a moral (as opposed to legal) right to someone to violate an ideal that, by Audi’s own reckoning, is “involuntary” in the sense that its “non-fulfillment [under the relevant conditions] subjects citizens in a liberal democracy to criticism” (85).
for, let alone attain, excellence in their role of citizens? Can we \textit{demand} excellence of our fellow citizens? Surely, to excel in some activity is to be more than just good: to excel is to be good to a very high degree, to “surpass or do better than others,”\textsuperscript{31} to be set apart from one’s peers and possibly win their admiration for doing exceptionally well compared with the average performance.

Of course, we could just stipulate that an ideal of citizenship is something only likely to be attained by the few who excel in their activities as citizens, and invite people to strive to imitate these rare examples, but not suggest that failure to meet the ideal is something blameworthy or to be condemned. If this is what we mean by an ideal, then an ideal is “superogatory” or morally commendable but not considered morally obligatory. In other words, we praise those who meet the ideal but we do not \textit{blame} or view as guilty of serious moral failure those who fail to meet it. For example, we praise citizens who volunteer to clean up their neighborhood but we do not necessarily \textit{blame} or view as morally reprehensible citizens who stay at home.

Rawls occasionally speaks of his ideal of citizenship as if this were the sort of thing he had in mind: a superogatory ideal that citizens may voluntarily embrace but that is not morally required or considered morally obligatory.\textsuperscript{32} However, there are two difficulties with this interpretation: first, it is hard to see how a superogatory ideal that does not impose moral \textit{requirements} upon citizens can penetrate the dealings of citizens

\textsuperscript{31} See \textit{Webster’s II New College Dictionary}.

\textsuperscript{32} First, Rawls consistently describes his normative picture of citizenship as an “ideal,” a word many more naturally associate with the aspirational than the obligatory (e.g. PL, 216-218). Secondly, Rawls repeatedly insists that citizens embrace the “ideal of public reason” (and presumably any concomitant ideal of citizenship) not simply on its own terms, but “from within their own reasonable doctrines” (PL, 218). This would seem to suggest that it is only binding insofar as it is sanctioned by citizens’ comprehensive doctrines, or views of the good.
and structure their behavior to the extent necessary in order for the ideal to purchase the stability and mutual respect and trust that Rawls is attempting to secure for the liberal polity. If Rawls’s ideal of citizenship were really an optional aspiration that citizens could evade without social sanction or blame, it could not serve as it is meant to, as the glue of social order.\(^{33}\) Second, the Rawlsian ideal imposes moral duties upon citizens\(^{34}\), and it is hard to conceive of duties as optional aspirations or excellences whose non-attainment occasions regret rather than moral disapproval.

I have argued that there are aspects of Rawls’s ideal of citizenship, regarding both its content and its function in his theory, that counsel against reducing it to supererogatory virtue. We have seen, for example, that Rawls’s ideal imposes duties upon citizens. Should we, then, consider an ideal in this context as a set of civic duties or moral requirements associated with citizenship? Although it may at first seem strange to apply the concept of an ideal to a set of duties, once we notice that a pattern of excellences may be partial or incomplete rather than comprehensive, there is no reason to refuse to consider the disposition to fulfill a set of duties as an aspect of excellence in citizenship. True enough, such duties seem unlikely to exhaust the field of excellence in citizenship, but Rawls never claims that his ideal of citizenship is exhaustive or complete in this sense.

\(^{33}\) One of the basic conditions of political stability is that “people who grow up under just institutions (as the political conception defines them) acquire a normally sufficient sense of justice so that they generally comply with those institutions” (PL, 141). The Rawlsian ideal of citizenship is designed to meet this condition by proposing a model of citizenship capable of grounding a “reasonable moral psychology” conducive to a reliable sense of justice (PL, 86).

\(^{34}\) See, for example, PL, 217: “the ideal of citizenship imposes a moral, not a legal, duty—the duty of civility—to be able to explain to one another on [certain] fundamental questions how the principles and policies they advocate and vote for can be supported by the political values of public reason.”
If we had to choose between a superogatory and a deontological reading of the Rawlsian ideal of citizenship, I would choose the deontological reading. However, this is too crude a choice, and already begs too many questions. For example, the disposition to be socially cooperative and “reasonable” in one’s dealings with others is clearly a good thing, a quality that is highly desirable in a citizen. But does a citizen have a “duty” to be socially cooperative and “reasonable” in every situation that seems to call for such dispositions? Could there be situations in which it is unusually difficult to exercise these virtues and therefore in which we do not hold a person morally blameworthy or reprehensible for failing to do so?35 Since Rawls maintains the language of “ideals”36 but does not resolve that language into deontological or superogatory/perfectionist categories, and since such a resolution would appear to create unnecessary difficulties for his ideal, I too will refrain from attempting such a resolution, and instead allow that an ideal may have both superogatory and deontological connotations, depending on the context.

Assuming my analysis of the concept is correct, it would seem that the definition of “ideal” most in keeping with Rawls’s project is, roughly, a “morally desirable pattern of action and dispositions.”37 The proposed definition leaves open the ultimate grounds for considering any particular element of the pattern “morally desirable” and does not pre-empt arguments over the precise normative status of the pattern as a whole (beyond

---

35 Imagine, for example, if one were repeatedly taunted by one’s political adversaries. Would it be “unreasonable” to cast civility aside and verbally attack one’s adversaries? Perhaps, but would there be a duty to forbear from “fighting fire with fire”? It seems to me that a plausible argument could be made in either direction.

36 Rawls’s choice of the word “ideal” is not an isolated one: he consistently uses the language of ideals, both in reference to public reason, and in reference to citizenship. See, for example, PL, 84, 213, 216-218, 253.

37 The use of the adverb “morally” should not here be taken to imply any special, e.g. Kantian, understanding of the “moral” sphere as opposed to, say, the ethical or aesthetic spheres. By “morally desirable” here, I simply mean, shown to be choiceworthy to a rational agent by good reasons (for action).
supposing that the pattern is to be desired or endorsed on moral grounds). Consequently, it is likely to be acceptable to most theorists, whether of classical/Aristotelian or modern orientation.\textsuperscript{38} In short, this definition of “ideal” conveniently captures the terrain of the discussion in neutral terms or terms acceptable to the various participants.\textsuperscript{39}

There is one more possible connotation of the term “ideal” that is worth exploring: if an ideal is a “morally desirable pattern of action and dispositions,” the feasibility of the ideal, that is, when and where it can and should be adhered to, remains an open question. Of course, we might imagine an ideal which commanded us to adhere to it even if “hell freezes over.” Perhaps aspects of our ideals may function in this way, e.g. the moral prohibition of certain acts such as lying under oath, torture, and killing for pleasure, the precept of charity to share out of our abundance with those in need, and so on. However, ideals or recognizable “patterns of excellence” tend to take on the color of their cultural and historical surroundings, and often presuppose the existence of certain social and political conditions. For example, the ideal of statesmanship in a large-scale industrialized society will likely impose different demands regarding contact with constituents than an ideal of statesmanship in the relatively small-scale Athenian polis, even if some principles such as accountability and honesty are present in both ideals. Similarly, ideals of tolerance and freedom in a liberal democracy presuppose certain favorable cultural and political conditions, and absent those conditions (perhaps an obvious example is a minimal degree of public order and political stability), such ideals

\textsuperscript{38} For example, arguably the very distinction between “deontological” and “supererogatory” virtue is meaningless or at best highly contrived from a classical Aristotelian perspective.

\textsuperscript{39} The main position I engage here will be Rawls’s, and for my immediate purposes, what matters is that the definition of “ideal” presents the problem in a way that is neutral between Rawls’s and my view of the truest or most apt ideal of citizenship for a liberal democracy.
move from the zone of immediately binding precepts or norms to long-range aspirations.\textsuperscript{40} Of course, we may disagree over the precise conditions under which a given ideal is no longer immediately practicable or relevant, but that such conditions exist is hardly disputed at all.

Of course, strictly speaking we may imagine an ideal that is unconditionally binding either because it is formulated in sufficiently general terms to allow its adaptation to particular situations and socio-cultural contexts, or because the ideal itself already contains the relevant feasibility constraints. The latter possibility does not seem plausible. After all, it seems highly unlikely that even the most elaborate and complex ideal could include within itself an exhaustive list of the conditions under which it applies and does not apply. There are too many potential defeaters or qualifiers of an ideal, some of which may be associated with unanticipated future developments such as revolutionary technological developments, to be exhaustively enumerated with any confidence. As far as more general ideals are concerned, on the other hand, it is certainly arguable (though clearly disputable as well) that ideals such as the Christian one of “love your neighbor as yourself” are sufficiently general that their spirit may be preserved in their application to particular cases, even those cases that appear to challenge the practicability of the ideal.

Unconditional ideals, interesting and conceivable as they are, are not my main concern here. Rather, I am concerned with ideals of citizenship. Now, although a plausible case can be made for the possibility of an ideal that is not altered in its essential

\textsuperscript{40}This is reminiscent of Hobbes’s distinction between precepts that bind \textit{in foro interno} and those that bind \textit{in foro externo}. See Thomas Hobbes, \textit{Leviathan}, ed. Edwin Curley (Indianapolis and Cambridge: Hackett Publishing Company, 1994) [1668], 99 (ch. xv): “The laws of nature oblige \textit{in foro interno}, that is to say, they bind to a desire they should take place; but \textit{in foro externo}, that is, to the putting them in act, not always. For he that should be modest and tractable, and perform all he promises, in such time and place where no man else should do so, should but make himself a prey to others, and procure his own certain ruin, contrary to the ground of all laws of nature, which tend to nature’s preservation.”
demands by social and historical conditions, no such case can be made for ideals of
citizenship, for two reasons in particular: first, ideals of citizenship concern a role that is
structured, at least to some extent, by contingent features of a society such as its mode of
government, its forms of civility, and its laws; secondly, ideals of citizenship are not just
philosophical abstractions to be studied and developed by intellectuals, but images or
exemplars to be grasped and imitated by citizens from a wide array of social and
educational backgrounds. Therefore, their requirements must of necessity be relatively
concrete and unencumbered by qualifications, and they must be already tailored to the
society in question. This is not to suggest that ideals of citizenship provide readymade
answers to the practical question, “what is to be done here and now?” But it is to suggest
that the guiding values and principles embedded in such ideals are inevitably conditioned,
at least in part, by social and historical contingencies.

Of course, for those who expect moral ideals to be pure Platonic forms unaffected
by the contingencies of history, this discussion may seem not only dissatisfying but
subversive of morality itself. In response, I should say a few things to distinguish my
view of ideals of citizenship from cultural relativism or historicism. The best way to
explain my position is to elucidate the relation between practical ideals and practical
reason more generally. Practical ideals provide us with a picture of morally desirable
action and dispositions. Such ideals may be highly personalized, say, my view of what
sort of life I in particular want to achieve (e.g. I want to have an enriching relationship
with so-and-so; I want to express my religious faith in such-and-such a way of life); or
they may be relatively impersonal, providing general standards of action that can be
readily grasped and implemented by people from all walks of life. Among the latter, we
find Christianity, Buddhism, secular humanism, Communism, and various liberal democratic creeds.⁴¹

Now, just as practical reasoning is guided by certain general principles, it is aided by ideals of various sorts, which are to be distinguished from discrete principles or precepts since unlike the latter, ideals consist in desirable patterns of action and dispositions clustered around particular life-goals or ways of life thought to be desirable all-round. Moral or ethical ideals are contingent in two ways. First, they vary across space and time—after all, it only stands to reason that what I consider a desirable standard of behavior or way of life may change in significant ways if my cultural and social environment changes significantly. For example, Aristotle could not have had anything closely resembling a modern liberal democratic ideal, not only because of the limits of his moral insights (reflected, for example, in his acceptance of slavery⁴²), but because such an ideal does not speak to the social structure and moral demands of ancient Greece.⁴³ Furthermore, every age and society has its own distinctive challenges and its

⁴¹ Of course, even these ideals must be given particular expressions and so we can distinguish between the general ideal and its specific implications for me, for my life or for this group of people.

⁴² Of course, to say that Aristotle’s views on slavery were wrong, mistaken, or less enlightened than ours is not to blame him for holding those mistaken views. It could be argued, for example, that slavery was so central to ancient Athenian culture and the Athenian way of life that even the most astute philosopher’s acceptance of it was understandable if not entirely excusable.

⁴³ For example, the private-public distinction had not yet been drawn to assign special dignity to the private sphere and thus many people valued public goods such as political participation, honor, and rhetorical excellence more highly than private goods such as family relations, while practices now considered by many to pertain to the "private sphere," such as religion, were counted as public and even political practices by ancient Greeks (for an illuminating and provocative discussion of the contrast between ancient and modern understandings of the private-public distinction, see Hannah Arendt, The Human Condition, 2nd ed. (London and Chicago: The University of Chicago Press, 1998); Benjamin Constant, “The Liberty of the Ancients Compared with That of the Moderns: Speech Given at the Athene Royal in Paris,” in Political Writings, ed. Biancamaria Fontana (Cambridge: 1988)). These differences would have given rise to a different structure of values and desires, so that the “liberties of the moderns” would have meant little to Greek citizens, and any attempt to implement the liberal democratic ideal would have been tantamount to “forcing men to be free.”
own distinctive social conventions or stylized forms of politeness and interaction, which inevitably inform its moral or ethical ideals.\footnote{For example, in Britain up to quite recently it was considered disrespectful to address a stranger, e.g. a new client, by his or her first name; whereas in the United States, it is increasingly considered a sign of good will and \textit{bonhomie} to address a stranger, e.g. a new client, by his or her first name.}

Secondly, ethical ideals are locally contingent in the sense that even within a single cultural setting, they express general standards to be applied with prudence rather than categorical imperatives pure and simple. Of course, ideals may embody unconditional norms—this is not disputed here—but to view all their requirements as unconditionally binding or as providing an exhaustive set of rules for behavior, is essentially to do away with the need for prudence.\footnote{For some interesting treatments of the modern neglect of the need for prudence in practical reasoning, see Julia Annas, “Being Virtuous and Doing the Right Thing,” delivered at Philosophy Colloquium at Notre Dame, 26 September 2003; Charles Larmore, \textit{Patterns of Moral Complexity} (Cambridge: Cambridge University Press, 1987), esp. chap. 1, “Moral Judgment – An Aristotelian Insight”; and Martin Benjamin, \textit{Splitting the Difference: Compromise and Integrity in Ethics and Politics} (Kansas: University Press of Kansas, 1990), chap. 5, “Judgement and the Art of Compromise,” esp. 107-121.} For example, although ideals of citizenship standardly consider law-abidingness a virtue, prudence is required to determine which laws one may obey in conscience consistent with remaining or growing as a good person all-round.

Now, it is important not to mistake the partial contingency of moral or ethical ideals for some brand of cultural or historical relativism. In pointing to the social and historical contingency of ideals, or the fact that they must adapt themselves to the culture and society at hand, I am not suggesting for a moment that moral ideals are \textit{entirely} conditioned by their social and historical location. Rather, I am suggesting that they are \textit{partially} shaped by it. Ideals and especially ideals concerning social roles such as that of citizen, are undeniably marked by their age and society in a variety of ways. However,
the fact that they are roles ascribable to *human persons* in *human societies* has a moral significance that transcends social and historical contingencies. For example, although the role of citizen in seventeenth century England, third century Athens BC, and Ciceronian Rome all differed in significant ways, we can identify certain qualities or virtues that have been prized in citizens wherever something resembling this role has existed, as well as vices that have been condemned. Thus, qualities such as loyalty, valor, compassion, magnanimity, generosity, law-abidingness and truthfulness have been prized in citizens throughout history, though interpretations of such virtues have no doubt diverged in various ways. Similarly, betrayal, cowardliness, dishonesty, selfishness, pettiness and contempt for the law have been consistently condemned in citizens, even if the interpretations of such vices have not always perfectly converged.46

What I would like to suggest is that it is no coincidence that we can identify certain core virtues of citizenship that have been recognized, to the best of our knowledge, as long as the role has existed. Since citizenship is a role that defines the relations among human persons participating in a shared political and social order, and since paradigms of human virtue and vice have remained remarkably stable over time, it is hardly surprising that patterns of civic excellence, though varying across regimes and across history, do reveal certain underlying virtues or excellences that are invariant across time, even if their perceived requirements do vary. The most obvious explanation for such stability across time and culture is the thought that human beings, wherever and

46 Obviously, I am not arguing that *everyone* who has thought about or written about citizenship has adhered to this paradigm, but nonetheless, exceptions aside, this has been the dominant way of thinking about citizenship. For an interesting treatment of the origin and development of the concept and praxis of citizenship in ancient Athens, see Philip Brook Manville, *The Origins of Citizenship in Ancient Athens* (Princeton, N.J.: Princeton University Press, 1997 [1990]).
whenever they are born, are at their core the same sort of being, a being with certain fundamental moral capacities such as the capacity to reason and make choices, and the ability to enter into meaningful and loving relationships with other human beings.\footnote{Of course, many human beings do not in fact possess such capacities either because they are undeveloped (e.g. a baby) or because they have been irreparably damaged (e.g. brain damage). This raises the question of the moral status of humans lacking what we take to be those capacities especially distinctive of humans. However, I am only considering here ideals of citizenship, which necessarily presuppose the possession of the relevant capacities at least to some degree (this is why, for instance, infants cannot be thought of as exemplars of good citizenship).}

Whether we call this human nature, or the fact of being human, something along these lines seems the best explanation of the transcultural and transhistorical dimension of ethical and moral life.\footnote{This dissertation is not intended as a defense of moral objectivity or of the enduring moral significance of human nature. However, for a defense of morality that gives pride of place to human nature, I would refer the reader to such works as John Finnis, *Natural Law and Natural Rights* (Oxford and New York: Oxford University Press, 1980); H.L.A. Hart, *The Concept of Law*, 2nd ed. (Oxford: Oxford University Press, 1994 [1961]), esp. chap. 9, “Laws and Morals”; Fulvio di Blasi, *God and the Natural Law*, trans. David Thunder (St. Augustine Press, forthcoming); Russell Hittinger, *The First Grace: Rediscovering the Natural Law in a Post-Christian World* (Wilmington, Delaware: ISI Books, 2003).}

To sum up, an ideal of citizenship can be thought of as (a) a pattern of morally desirable action and dispositions pertaining to the special role of citizen, (b) potentially having superogatory as well as deontological connotations. Finally, (c) the ideal is conditioned by both contingent factors such as social and political structures and conventions, and invariant factors such as human nature or our humanity, insofar as this holds certain fundamental values constant over space and time. I should stress that my account of ideals of citizenship goes well beyond anything Rawls is committed to, but it does serve to situate my interpretation of his ideal and the argument over ideals within a broader moral context, and it also serves to prepare the ground for my later discussion of the relation between integrity and the role of citizenship (see especially §6.1 and 6.2.3).
§2.2 The Theoretical Context

Rawls’s ideal of citizenship cannot be understood independently of the larger goals of his theory and the intended contribution of his ideal of citizenship to these larger goals. The overarching goal of Rawls’s theory of politics is to provide an acceptable basis for citizens of diverse worldviews and moral perspectives to live peacefully together.49 This basis can be specified by three conditions:

(i) First, social and political order must be based on moral principles rather than on a mere “modus vivendi,” viz. a pragmatic agreement based on shifting coalitions of individual and group interests.50

(ii) Second, the fundamental basis of political and social order, what Rawls calls the “constitutional essentials and matters of basic justice,” must be secured through the reasoned rather than coerced agreement of citizens.51

(iii) Third, the political regime must win a sufficient degree of support among citizens to be stable or provide a reliable basis for public order.52

In other words, we can best make sense of Rawls’s entire project if we keep in mind the overarching goal of providing an acceptable basis for citizens to live peacefully together.

---

49 Thus, the fundamental question driving Rawls’s project, as stated in PL, is “How is it possible that there may exist over time a stable and just society of free and equal citizens profoundly divided by reasonable though incompatible religious, philosophical, and moral doctrines?” (PL, xx).

50 A modus vivendi, as Rawls defines it, is a “social consensus founded on self- or group interests, or on the outcome of political bargaining: social unity is only apparent, as its stability is contingent on circumstances remaining such as not to upset the fortunate convergence of interests” (PL, 147).

51 This is implied by the “liberal principle of legitimacy,” which states that “political power is legitimate only when it is exercised in accordance with a constitution…the essentials of which all citizens, as reasonable and rational, can endorse in the light of their common human reason” (John Rawls, Justice as Fairness: A Restatement, ed. Erin Kelly (Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press, 2001), 41).

52 As Rawls puts it, “an enduring and secure democratic regime, one not divided into contending doctrinal confessions and hostile social classes, must be willingly and freely supported by at least a substantial majority of its politically active citizens” (PL, 38).
together in a pluralistic society, which is (i) *principled, not merely pragmatic*; (ii) based on *reasoned, not coerced agreement*; and (iii) *sufficiently supported by citizens to be stable, not precarious*. Keeping in mind this complex goal of Rawls’s theory, I offer here a rough overview of political liberalism, a convenient label I will use—following Rawls and his interpreters—as shorthand for Rawls’s conception/theory of the well-ordered polity. Each element of Rawls’s theory can be seen to contribute (if not in fact at least by design) to a viable regime as specified by the three conditions just enumerated. The most important elements of political liberalism are the “fact of reasonable pluralism”; the concept of the citizen as free and equal; the concept of society as a fair and stable system of social cooperation enduring over time; the distinction between a “political conception” and a “comprehensive doctrine”; the distinction between public and nonpublic reason; and the idea of a well-ordered society. Rawls’s ideal of citizenship is so intimately bound up with these concepts that any attempt to treat them separately from the ideal of citizenship will appear artificial or contrived. However, it is analytically helpful to focus on the overall theory and the ideal of citizenship in successive stages, if for no other reason than that this will underscore how substantial and complex the philosophical foundation of the Rawlsian ideal of citizenship is, and how inseparable the ideal is from the wider theory of political liberalism.

Before proceeding to lay out the basic elements of Rawls’s theory, I should note a discrepancy between my order of exposition and Rawls’s. First of all, Rawls does not include the fact of reasonable pluralism among the “fundamental ideas” of political liberalism (PL, Lecture I), while I do. I do not believe this suggests any fundamental disagreement, but a matter of emphasis. The “fact of reasonable pluralism” is one of the
background assumptions throughout political liberalism, so much so that at times it can be almost invisible. I make it explicit as a fundamental idea of political liberalism, whereas Rawls seems to treat it more as a basic presupposition that can be largely taken for granted as a description of our society. Secondly, Rawls does not list the idea of public reason as one of the fundamental ideas of political liberalism. Given the pivotal role it plays in his conception of a stable and just political order, this may seem surprising at first. However, I would surmise that he views public reason as largely derivative from other ideas such as the idea of society as a fair system of social cooperation, and the idea of citizens as free and equal. It is sufficiently central to political liberalism, however, that I feel justified in treating it as a fundamental idea.

Thirdly, I do not include the original position as a fundamental idea of political liberalism, partly because I see it as implicit in the idea of society as a fair system of cooperation (insofar as the original position is a conceptual device for yielding the principles of a fair socio-political order); and partly because I do not see the original position, which is, after all, a particular device of representation used to model the principles of justice, as a necessary presupposition of political liberalism—it has derivative status with respect to, say, the idea of social cooperation or the idea of the citizen as free and equal. In other words, it is not itself of fundamental importance: rather, it “models what we regard—here and now—as fair conditions under which the representatives of free and equal citizens are to specify the terms of social cooperation in the case of the basic structure of society; and…it also models what…we regard as acceptable restrictions on reasons available to parties for favoring one political
conception of justice over another…” The derivative status of the original position within political liberalism as a whole is also suggested by Rawls’s comment that “I have proposed that one way to identify [reasonable] political principles and guidelines is to show that they would be agreed to in what in *Political Liberalism* is called the original position.”

Fourthly and finally, Rawls’s order of exposition is quite different to mine. It goes as follows: the idea of a political conception of justice; society as a fair system of cooperation; the original position; the political conception of the person; and the well-ordered society. Whereas mine proceeds as follows: the fact of reasonable pluralism; the political conception of the person; society as a fair system of social cooperation; the idea of a political conception of justice; public reason; and the well-ordered society. The main discrepancy to be noted is that I treat the political conception of the person much earlier than Rawls; and prior to the idea of a political conception of justice. How can I justify this discrepancy? The main rationale behind it is that I do not see how the idea of a political conception of justice can be motivated without assuming the importance of respecting the freedom and equality of citizens, under some description. In other words, I do not see why a political philosopher would even frame the distinction between a political conception and a comprehensive doctrine unless there were already a concern to offer citizens terms of cooperation they could reasonably accept, as free and equal. By preceding the idea of a political conception of justice with the political conception of the person, I do not mean to suggest that the “political” conception of the person need be

---


viewed as comprehensive or pre-political; but I do mean to suggest that its “political” or “non-comprehensive” character is only self-consciously conferred “after the fact.” That is to say, it is the conception of the person as free and equal that motivates the idea of a political conception of justice, and it is only subsequently that the conception of the person as free an equal can be interpreted as “political, not metaphysical.” In other words, I reject the notion that the political conception of the person just “falls out of” the political conception of justice, viewed as its ground. On the contrary, I see the political conception of the person as providing an essential raison d'être to the political conception of justice.

§2.2.1 The Fact of Reasonable Pluralism

The starting point for political liberalism is the “fact of reasonable pluralism.” People come to widely differing conclusions not only about which objects, actions or ways of life are valuable all things considered, but what the fundamental measures of value are. In other words, moral disagreement runs deep in modern societies. But add to this the observation that there are many plausible explanations for disagreement besides malice, carelessness, stupidity, selfishness, irrationality, etc.—explanations compatible with the sincerity, conscientiousness, and responsibility of parties on both sides of the disagreement—and we have the “fact of reasonable pluralism”: the fact that many reasonable people can and do reasonably come to widely differing and incompatible moral conclusions. The factors motivating “reasonable disagreement” over values are what Rawls calls the “burdens of judgment”: the burdens of judgment are factors that differentially affect the moral reasoning of agents without rendering them vicious or
defective in any obvious way. Rawls lists six of the more obvious factors but there are probably many others:

(1) The empirical evidence bearing on a case is conflicting and complex, and thus hard to assess.

(2) Different considerations, even assuming their relevance is accepted by all, may be weighted differently by different persons.

(3) Concepts are often vague and indeterminate with respect to hard cases.

(4) Our moral and political judgments are influenced by our total life experience, which necessarily differs across persons.

(5) Normative considerations may be advanced in favor of both sides of an issue, producing a conflict of values difficult to mediate.

(6) A hard choice between many possible goods must be made in social and political policy-making. There is a limited social space which cannot instantiate the full range of values.\(^{55}\)

Now, the “fact of reasonable pluralism” is not a claim about human nature as such, but a claim about the conditions that distinguish moral inquiry in late modern post-industrial societies such as those of Western Europe and North America. That \textit{some degree} of “reasonable pluralism” could have existed in ancient or medieval societies is not denied, but the “fact of reasonable pluralism,” insofar as it sets up a political problem, indicates the presence of significant moral disagreement of this nature on a historically unprecedented scale, with significant and historically unprecedented consequences for political order.

The fact of reasonable pluralism, taken by itself, is of course insufficient to yield Rawls’s conclusion that we ought to attempt to abstract from deep-seated moral disputes among reasonable citizens when forging the terms of political association. One can well imagine somebody accepting the “fact of reasonable pluralism,” but still believing that the moral truth on this or that policy is too important to await the agreement of all or most “reasonable citizens”: such a person may do everything in her power to implement the moral truth as she sees it even in the face of the “reasonable” dissent of her fellow citizens. Their dissent, our reformer may admit, is reasonable in the sense that it is not based on irrationality or bad faith; but even reasonable dissent may be overridden or disregarded if the moral cause is important enough.\(^{56}\) I give this example not as an argument for overriding reasonable dissent but to show that something besides the fact of reasonable pluralism is required to explain the liberal’s reluctance to coerce citizens in such cases. That something is the concept of the citizen as free and equal.

§2.2.2. The Concept of the Citizen as Free and Equal

Nothing drives the engine of political liberalism as much as Rawls’s conception of citizens as free and equal persons. The first thing to be said about this conception is that no political philosopher has a monopoly on the meaning of freedom and equality—on the contrary, these are what William Connolly would describe as “essentially contestable” concepts in political theory.\(^ {57}\) Consequently, it is of course a particular

---

\(^{56}\) The obvious example is opposition to abortion, which is viewed by many as a sufficiently grave human rights abuse to justify overriding reasonable dissent to anti-abortion laws.

interpretation of citizens as free and equal that informs political liberalism, and it is this interpretation, rather than freedom and equality as such, that interests us here.

For Rawls, the equality of citizens consists in the fact that they are each entitled to the same fundamental human and civic rights and liberties as the next person; the interests of each are to be given equal consideration to the extent practicable in the design and day-to-day operation of political institutions. As Rawls puts it, citizens are to be viewed by political authorities as “self-authenticating sources of valid claims.” In other words, the bare fact that a claim is advanced by a citizen already makes that claim prima facie worthy of consideration, at least if it does not obviously violate the requirements of the public conception of justice. Dworkin’s formulation though not identical to Rawls’s, is, I believe, faithful to Rawls’s intentions: each citizen is due the “equal concern and respect” of government and its agents. The distinctiveness of this notion of “equal concern and respect” is evinced by contrasting it with the utilitarian notion that the happiness of any given citizen can be effectively dispensed with if doing so adds more overall than alternative courses of action to the total happiness or utility of all individuals.

---

58 See PL, 32-33: “A second respect in which citizens view themselves as free is that they regard themselves as self-authenticating sources of valid claims. That is, they regard themselves as being entitled to make claims on their institutions so as to advance their conceptions of the good (provided these conceptions fall within the range permitted by the public conception of justice). These claims citizens regard as having weight of their own apart from being derived from duties and obligations specified by a political conception of justice, for example, from duties and obligations owed to society. Claims that citizens regard as founded on duties and obligations based on their conception of the good and the moral doctrine they affirm in their own life are also, for our purposes here, to be counted as self-authenticating. Doing this is reasonable in a political conception of justice for a constitutional democracy, for provided the conceptions of the good and the moral doctrines citizens affirm are compatible with the public conception of justice, these duties and obligations are self-authenticating from a political point of view.”

considered as a collective. Indeed, Rawls himself goes to great lengths to separate his understanding of equality from that of utilitarians.\textsuperscript{60}

The equality of citizens is of course reflected in Rawls’s first principle of justice, which states that citizens are each entitled to a fully adequate scheme of basic liberties\textsuperscript{61}; and in the second principle, which mandates that material inequalities, insofar as they are permitted, must be to the maximal advantage of the least well-off.\textsuperscript{62} But how is the fundamental concept of equality that lies behind the two principles of justice to be given meaningful content? I believe the Rawlsian conception of equality can be summed up in the claim that citizens have an equal entitlement to the conditions necessary for the adequate development of their two basic moral powers, namely, the capacity for a sense of justice; and the capacity for a conception of the good. The development of these two powers constitutes the moral autonomy of citizens. So, to restate Rawls’s conception of equality: Citizens are equal insofar as they are equally entitled to enjoy the basic social and economic conditions necessary for moral autonomy.

But what is autonomy, and what are its conditions? Let us start by considering the basic constituents of autonomy: first, citizens have “the capacity to form, to revise, and

\begin{itemize}
\item \textsuperscript{60} See, for example, John Rawls, \textit{A Theory of Justice}, 2nd ed. (Cambridge, MA: The Belknap Press of Harvard University Press, 1999 [1971]) [hereafter TJ], 24-30; PL, 179-180.
\item \textsuperscript{61} This principle underwent some reworking under the pressure of Hart’s critique, as acknowledged by Rawls himself. In the original formulation, the first principle read “Each person has an equal right to the most extensive total system of equal liberties compatible with a similar system of liberty for all” (TJ, 220); whereas the revised formulation reads, “Each person is to have a fully adequate scheme of equal basic liberties which is compatible with a similar scheme of liberties for all” (PL, 291). For a discussion of the revision, see PL, 289-294. Cf. H.L.A. Hart, "Review of Theory of Justice," \textit{University of Chicago Law Review} 40 (Spring 1973): 535-555.
\item \textsuperscript{62} The two principles, as stated in PL, read as follows: “a. Each person has an equal right to a fully adequate scheme of basic liberties which is compatible with a similar scheme of liberties for all. b. Social and economic inequalities are to satisfy two conditions. First, they must be attached to offices and positions open to all under conditions of fair equality of opportunity; and second, they must be to the greatest benefit of the least advantaged members of society” (291).
\end{itemize}
rationally to pursue a conception of one’s rational advantage or good”; second, citizens have the capacity to “understand, to apply, and to act from the public conception of justice which characterizes the fair terms of social cooperation.” Rawls calls these the two “moral powers” of persons. Just insofar as persons possess these powers “to the requisite minimum degree,” they may be viewed as political equals. The two moral powers are reasonably viewed by the political theorist or constitution-designer as the core constituents of a meaningful human life, and Rawls insists that citizens have what he calls a “higher-order interest” in developing these powers. This seems to suggest that these are interests that can be imputed to us in virtue of certain features of our humanity—however culturally or socially conditioned they may be—rather than in virtue of our choices. They are interests that are not contingent upon our idiosyncratic tastes or propensities: they are not interests in the same region as, say, my interest in soccer or philosophy or cinema. We need not assume that they are absolute, unshakable interests, or interests entirely immune to our choices—some would argue, for example, that agents may legitimately prefer to die now for a noble cause than live on for years to develop their capacity for a sense of justice and a conception of the good—but they are more foundational, enduring, and universally shared than many other interests and to that extent seem to deserve a special category of their own—for Rawls, the category of the “moral powers.” Although the Rawls of Theory of Justice grounded the importance of the

---

63 PL, 19.

64 See PL, 74: “Since citizens are regarded as having the two moral powers, we ascribe to them two corresponding higher-order interests in developing and exercising these powers. To say that these interests are “higher-order” interests means that, as the fundamental idea of the person is specified, these interests are viewed as basic and hence as normally regulative and effective. Someone who has not developed and cannot exercise the moral powers to the minimum requisite degree cannot be a normal and fully cooperating member of society over a complete life.”
moral powers in the intrinsic worth of a rational, autonomous life, this avenue is closed to him in Political Liberalism, which must insist, in order to remain “political, not metaphysical” in orientation,\textsuperscript{65} that “the good involved in the exercise of the moral powers and in the public recognition of persons’ status as citizens belongs to the political good of a well-ordered society and not [at least, not necessarily] that of a comprehensive religious, philosophical, or moral doctrine.”\textsuperscript{66} We might say, then, that the significance of the two moral powers is derived, at least for political purposes, from the fact that they are almost universally recognized and experienced as “good” and of fundamental importance by persons within the political culture of constitutional democracies, rather than from any exalted view of rational autonomy.\textsuperscript{67}

Let us now consider more in depth the two moral powers we have a “higher-order” interest in developing. Consider first the capacity to form, revise, and rationally pursue a conception of the good: what exactly does Rawls have in mind here? First, a conception of the good is some view not only about one’s rational advantage, narrowly understood, but about “what is valuable in human life” more generally.\textsuperscript{68} Rawls is not suggesting that most citizens have tightly structured, well thought through plans of life—though of course some may. Nor is he suggesting that we have complete control over the content of our conceptions of the good, such that we can stand back from them entirely and choose them from some impartial-observer perspective. On the contrary, Rawls

\textsuperscript{65} See §2.2.4 below for an explanation of the “political” character of Rawls’s project.

\textsuperscript{66} PL, 203.

\textsuperscript{67} See PL, 202-203. That is to say, the equality and freedom of citizens, though it does amount to a form of rational autonomy, is not based on an independent moral ideal of rational autonomy derived from some comprehensive doctrine such as neo-Kantian moral theory.

\textsuperscript{68} PL, 19.
admits that citizens often have very loosely structured, shifting conceptions of their good, and that such conceptions are deeply informed by factors beyond their own control such as their upbringing and social ties. Nonetheless, people are capable of altering their views of the good bit by bit, and even of taking decisive steps to repudiate old conceptions and adopt new ones. They may find it impossible to conceive of life without certain fundamental purposes or constituent goods, but nonetheless they at least have enough dominion over their life that they can consciously embrace those purposes and goods, or distance themselves from certain purposes and goods, as the case may be. A meaningful human life is hard to imagine—at least on Rawls’s account—without a developed capacity to form, pursue, and revise a conception of the good.

The second of the two basic moral powers, the capacity to form and act upon a sense of justice, is essentially the capacity to regulate my pursuit of the good according to the requirements of life in society, in a way that is responsive to the interests of others. We not only have the capacity to (partially) determine the direction our own lives will take (what Rawls calls “rational autonomy”), but we also have the capacity to regulate our lives according to rules that are acceptable to others (what Rawls calls “full

---

69 According to Rawls, “many citizens may not hold any well-articulated comprehensive doctrine at all. Perhaps most do not. Rather, they affirm various religious and philosophical, associational and personal values together with the political values expressed by the political conception” (Rawls, Justice as Fairness: A Restatement, 33).

70 Sandel, in Michael Sandel, Liberalism and the Limits of Justice (Cambridge, London, New York: Cambridge University Press, 1982), charged Rawls with advancing a naïvely abstract or ‘thinned out’ conception of the person, divorced from prior aims and attachments, but Rawls pointed out in response that this was a misunderstanding of the original position, which was meant as a conceptual device of representation, not a metaphysical description of persons (PL, 27-28). The independence of persons from their conceptions of the good is not to be understood as the ability to adopt and dispense with them at will, but the fact that for political purposes, a person’s identity should not be viewed as inevitably defined by a single conception of the good. This is explained in PL, 30-32.

71 One example given by Rawls is the conversion of St. Paul (PL, 31).
We can take into account the interests of our fellows and seek out fair terms of cooperation. Assuming that the social dimension of human life is of fundamental importance, it is hard to imagine a meaningful moral life that is not socially cooperative or oriented towards justice (barring, perhaps, cases such as a situation reflecting Hobbes’s “war of all against all”). Indeed, if one is unwilling to develop, or incapable of acting upon, a sense of justice, one is likely to eventually submit under force to the terms laid down by society, which is obviously restrictive of one’s capacity for moral self-direction or autonomy.

One might argue, of course, that it is possible to live a meaningful life without a sense of justice. Who is Rawls to say what is and is not a “morally meaningful” life? Why not consider the amoral, glory-seeking, life a meaningful one? The first thing to be said in response to this objection is that Rawls is not aiming to convince the moral skeptic or those who have turned their backs on the need for social cooperation and the consideration of others and their interests. Rawls is speaking to reasonable people, who at a minimum are people who are committed in principle to living with others on fair terms of cooperation. Thus, reasonable people are precisely those who are capable of a sense of justice, however minimal it may be. Rawls does not demonstrate that taking account of

---

72 See PL, 306, for the distinction between “rational” and “full” autonomy.

73 Rawls himself does not explicitly connect the sense of justice with a morally meaningful life, but I think it is implicit in his ascription of “full autonomy” only to those citizens who “freely accept the constraints of the reasonable, and [whose political life,] in so doing….reflects that conception of the person which takes as fundamental their capacity for social cooperation” (PL, 306). On this conception, it is clear that a person would somehow be morally stunted or undeveloped were he to lack the capacity for social cooperation.

74 There are of course more controversial formulations of the reasonable, but for now I will stick to the less problematic formulation. Though some of Rawls’s formulations can be accused of vicious circularity, I do not think this more commonplace formulation is circular in a problematic sense. The sense it conveys is that there are some people with whom we can have rational and reasonable conversations about political matters, and those are the people who have some grasp of, and appreciation for, the demands of justice and
the interests of one’s fellows is a necessary component of a morally meaningful human life. Rather, this seems to be a point he is confident most readers will not need any elaborate argument to see.\textsuperscript{75}

As I suggested earlier, on Rawls’s account citizens have an equal entitlement to the social and economic conditions necessary for the adequate development of their two moral powers. These conditions are captured under Rawls’s notion of “primary goods,” which, contrary to some interpretations, are not all purely instrumental, but inclusive of both instrumental \textit{and} intrinsic-instrumental goods. For example, self-respect or self-esteem is clearly valued both for its own sake (intrinsic good) and insofar as it enables a person to effectively pursue and act upon a conception of the good (instrumental good).\textsuperscript{76} Rawls never pretends to offer an exhaustive list of the primary goods, but he does offer a list of the more obvious and basic ones, which are to be secured by a just liberal regime to the extent practicable for all citizens. The theory of primary goods is supposed to be neutral among a host of different lifestyles and moral values. According to Rawls, citizens should agree without any significant strain on their commitments that the primary goods are a necessary (though hardly sufficient) condition for a meaningful human life, fairness, even if they disagree in their interpretation of such requirements. For a fuller discussion of the meaning of reasonableness in political liberalism, see \S\textsuperscript{3.2.2} below.

\textsuperscript{75} However, we might speculate that a morally autonomous life must be a life that is reconciled with the reasonable demands of society rather than in conflict with them, since a denial of the legitimate claims of society upon one’s actions would prevent one from ordering one’s actions freely within the laws of a reasonably just society. An unreasonable person or a person without a sense of justice would probably find the constraints of a just society upon his actions intolerable and thus he would only conform to the law out of fear, not out of rational assent.

\textsuperscript{76} See PL, 318: “Self-respect is rooted in our self-confidence as a fully cooperating member of society capable of pursuing a worthwhile conception of the good over a complete life. Thus self-respect presupposes the development and exercise of both moral powers and therefore an effective sense of justice. The importance of self-respect is that it provides a secure sense of our own value, a firm conviction that our determinate conception of the good is worth carrying out. Without self-respect nothing may seem worth doing, and if some things have value for us, we lack the will to pursue them.”

43
and as such, that the government is morally warranted and even obligated to ensure that all citizens are adequately supplied with such goods.

Of course, the purported neutrality of Rawls’s list of primary goods, but most of all, the role of the primary goods in determining the relations between citizens and between citizens and their government, has been widely criticized by political philosophers. This is not the place to canvass such objections, but it is worth pointing out, by way of example, that even someone who accepts with Rawls that people (a) possess the two moral powers; and (b) have a higher-order interest in their adequate development and exercise, may dispute the sizable role Rawls gives to government in securing this interest. Robert Nozick, for example, develops a very different concept of entitlement which, if true, would justify a much less ambitious role for government, yet his theory appears to be consistent with Rawls’s account of the two moral powers and their importance in the lives of citizens.77

According to Rawls, the primary goods fall under five basic headings:

a. basic rights and liberties, …given by a list;
b. freedom of movement and free choice of occupation against a background of diverse opportunities;
c. powers and prerogatives of offices and positions of responsibility in the political and economic institutions of the basic structure;
d. income and wealth; and finally,
e. the social bases of self-respect.78


78 The list is quoted exactly as it appears in PL, 181. A more detailed list can be found in PL, 308-309.
With this admittedly crude summary of Rawls’s conception of citizens as free and equal persons in hand, it is worth considering the question of foundations. Where does Rawls’s conception come from? How is it grounded? We cannot fully answer this question until we explicate in section §2.2.4 the distinction between political conceptions and comprehensive doctrines. But for now, suffice it to say that Rawls does not intend his conception of the person to be grounded in any comprehensive metaphysical view about the way people are or how they are situated in the universe. Rather, his view of the person is “political,” in at least three senses: first, it is designed specifically with political institutions and interactions in mind, and in this sense it has a limited domain; second, it is built up from notions thought to be implicit within the public political culture of a liberal democracy, in this case, that of the United States in particular, rather than from a view of the truth independent from our political culture; and third, efforts are made to avoid relying upon any highly contested philosophical principles or doctrines upon which reasonable disagreement is to be expected: in this sense, Rawls’s political conception of the person may be considered “free-standing” from comprehensive doctrines.\(^{79}\)

One might still wonder whether the moral status of citizens is something prior to and foundational with respect to the political conception of justice, or simply something that “falls out of” the political conception. I believe the most plausible position is that the moral status of citizens is both foundational within the argument and a central element of the political culture of constitutional democracies—as such, it constitutes appropriate “material” if you will, for a political conception of justice. The goal of a theory of justice should be both to draw on common elements of our shared political culture (advancing

\(^{79}\) The notion of a “comprehensive doctrine” will be explained in §2.2.4.
distinctive but plausible interpretations of them) and to make sure those elements square with our “considered moral judgments.” The authority of the notion of equality, for example, is not grounded in the fact that “we happen to believe it around here,” but in the fact that it is true. Unfortunately, Rawls tries to get by without relying on the notion of truth, and so it is not clear how the political conception can be grounded except in the fact that these happen to be our considered judgments. So the Rawlsian position appears to be, that the moral status of citizens, while serving a foundational role within the political conception of justice, derives its authority from the fact that it is implicit within our public culture, not from the actual moral status of citizens as such.  

Finally, a note on the relation between the Rawlsian conception of citizens as free and equal on the one hand, and the fact of reasonable disagreement on the other. I earlier remarked that the fact of reasonable disagreement, by itself, cannot explain the liberal’s reluctance to legislate a moral or political principle in the face of “reasonable disagreement” over its validity or legitimacy. However, once we combine the fact of reasonable pluralism (or disagreement) with Rawls’s concept of citizen as free and equal, the liberal aversion to coercive settlements of issues plagued by “reasonable disagreement” becomes understandable. For if it is indeed the case that it is one’s capacity to form and act upon a conception of the good that one accepts and identifies with, rather than the content of one’s conception of the good, that explains one’s moral status in the political sphere, then any coercively-enforced principle, just insofar as it requires unreasoned, unwilling conformity to an external conception of the good, is pro

tanto, at least *prima facie*, an attack on one’s moral status. Therefore, if coercion must occur (as it must in any viable polity), the principles animating it or justifying it must, in some fairly robust sense, be *accessible* or reasonably acceptable to those upon whom the coercive measures fall. This is not to say that the justification of coercion must win the *actual* acceptance of all; but it is to say that such a justification must, in principle, appeal to reasons that *could* reasonably be accepted by all. For now, I leave this intriguing ‘modal’ condition hanging, but will subject it to closer scrutiny in §3.1 below.

§2.2.3 The Concept of Society as a Fair and Stable System of Social Cooperation

Rawls counts as “fundamental ideas” within political liberalism both the political conception of the person (as citizen), and the concept of society as a fair and stable system of social cooperation enduring over time.\(^81\) Whereas the political conception of the person hones in on those features of persons that political justice ought to be especially guided by, the Rawlsian concept of society (as a fair system of social cooperation) highlights those features of human relations and shared life that a theory of justice for modern liberal democratic societies must take into account. Political liberalism assumes that social cooperation is a great good, and a good that all (reasonable) citizens are in principle committed to pursuing and protecting. Citizens are assumed to want a society in which social cooperation is possible and facilitated by social and political institutions. Cooperation is assumed to be sought on terms *fair to all*, but also on terms

---

\(^81\) The first six headings within Lecture I, “Fundamental Ideas,” nicely summarize what for Rawls are the fundamental building blocks of his theory: the ideas of a political conception of justice; society as a fair system of cooperation; the original position; the political conception of the person; and the well-ordered society. My exegesis of Rawls in §§2.2.1-2.2.6 is guided by but not confined to these ideas.
that prove stable and may be carried forward in their essentials by succeeding generations. The so-called original position is a device of representation designed to assist citizens and philosophers in coming up with general principles for a society whose goal is taken to be social cooperation on terms that are \textit{fair to all} and enduring over generations.

The concept of fairness as used by contractarian thinkers like Rawls is not susceptible to any simple definition. However, it is useful to distinguish it from justice, since fairness is more closely tied to procedural aspects of justice, such as openness and inclusiveness in processes of dispute resolution, while justice seems to be a category that is more outcome-oriented. Fairness is more closely tied to the \textit{process} by which decisions are reached than the \textit{outcome} of the process. Or perhaps to put it more perspicuously, fairness judges the outcome as a function of the process through which it was reached rather than on its merits considered independently from the process. Thus, we can speak of a decision procedure that is fair to all, and yet say without contradiction that the result, through some mistaken calculations or erroneous judgments of the participants, failed to live up to the demands of an ideal of justice. Some would argue, for example, that the decision of an un-biased umpire in a tennis game is \textit{fair}, insofar as procedures were followed in good faith and without favoritism, but somewhat \textit{unjust}, insofar as the umpire accidentally misjudged the play.

\footnote{Rawls points out in response to Sandel’s anthropological critique of the original position that the original position is no more than a conceptual device of representation. At no point is Rawls suggesting that the (imagined) inhabitants of the original position are in a situation that actually reflects the human condition. On this point, see Sandel, \textit{Liberalism and the Limits of Justice}, esp. chap. 1, “Justice and the Moral Subject,” and Rawls’s rejoinder in PL, 26ff.}
The concept of fairness, as it is reflected in the original position, is basically the notion that the philosophical decision procedure for generating the principles of justice does not arbitrarily favor the interests or advantage of any particular group or individual in society. But more positively, a fair social arrangement according to Rawls, is an arrangement in which the benefits and burdens of cooperation are fairly or equally distributed across participants. One of the reasons Rawls’s interpretation of fairness is so controversial is that while many people accept the first component of the view, the notion that no group or person should be arbitrarily advantaged or disadvantaged by the process; they reject Rawls’s inference that the benefits and burdens of cooperation must be equally distributed across participants; or else they view the benefits, along Nozickean lines, as formal opportunities rather than welfare. A just society, on Rawls’s view, is a society in which the basic structure and the principles animating it can be viewed as “fair” in the sense of not arbitrarily favoring or disadvantaging anyone, and entailing a close to equal distribution of benefits and burdens among participants. Hence, Rawls dubs his conception of justice “justice as fairness.”

The basis of social cooperation must be stable in the sense that it is not easily destabilized or overturned, e.g. by shifting coalitions of interests, whether in the present generation or in future generations. One necessary condition for stability is the

---

83 As Rawls puts it in PL, 16, “Fair terms of cooperation specify an idea of reciprocity: all who are engaged in cooperation and who do their part as the rules and procedures require, are to benefit in an appropriate way as assessed by a suitable benchmark of comparison...the benefits produced by everyone’s efforts are fairly distributed and shared from one generation to the next.”

84 See, for example, Nozick, Anarchy, State, and Utopia.

fundamental fairness of the system of social cooperation towards all participants, whether of present or future generations. If it bestows arbitrary or unfair advantages on any party, it is liable to generate envy or resentment on the part of the disadvantaged and will for this reason be unstable and ultimately, on Rawls’s account, unjust. The fairness of political liberalism is modeled in the original position, both in *Theory of Justice* and in *Political Liberalism*.

The original position is not a historical description of any actual society or even any possible society, but a conceptual device of representation, designed to represent more vividly our intuitions about justice as fairness.86 The original position is an imaginary procedure in which citizens or their representatives meet and come to an agreement on the basic principles to regulate their political society. It is designed to select principles of justice for a society conceived as a system of social cooperation enduring over time. Citizens are constrained in their reasoning by a “veil of ignorance” which robs them of knowledge of certain particulars about their own characteristics or likely position within society that are judged (a) likely to contaminate the choice of principles by allowing them to be designed to uniquely favor the position or preferences of this or that party to the choice; and (b) unnecessary and thus dispensable for the task of selecting principles of justice. The (imaginary) parties to the original position are rational in the sense that they are simply self-interested means-end reasoners, who are not already informed by a sense of justice. However, the constraints placed upon their reasoning (chiefly the veil of ignorance) are designed to ensure that the outcome is reasonable, in

---

86 This is stressed by Rawls in his response to Sandel’s metaphysical critique of the original position in *PL*, 25-27.
the sense that it is indeed consistent with the basic requirements of justice. Aspects of our knowledge that are screened out by the veil of ignorance include one’s place in society, his class position or social status...his fortune in the distribution of natural assets and abilities, his intelligence and strength...Nor again does anyone know his conception of the good, the particulars of his rational plan of life, or even the special features of his psychology such as his aversion to risk or liability to optimism or pessimism. More than this, I assume that the parties do not know the particular circumstances of their own society. That is, they do not know its economic or political situation, or the level of civilization and culture it has been able to achieve. The persons in the original position have no information as to which generation they belong.

Aspects of our knowledge that are let through include

[the fact] that [one's] society is subject to the circumstances of justice and whatever this implies. It is taken for granted, however, that [the parties] know the general facts about human society. They understand political affairs and the principles of economic theory; they know the basis of social organization and the laws of human psychology. Indeed, the parties are presumed to know whatever general facts affect the choice of the principles of justice. There are no limitations on general information, that is, on general laws and theories, since conceptions of justice must be adjusted to the characteristics of the systems of social cooperation which they are to regulate, and there is no reason to rule out these facts.

---

87 Rawls is very clear on the distinction between the rationality of the (imaginary) parties to the original position and the reasonableness of the outcome of their reasoning under the relevant constraints: “While the original position as a whole represents both moral powers, and therefore represents the full conception of the person, the parties as rationally autonomous representatives of persons in society represent only the rational: the parties agree to those principles which they believe are best for those they represent as seen from these persons’ conception of the good and their capacity to form, revise, and rationally to pursue such a conception, so far as the parties can know these things. The reasonable, or persons’ capacity for a sense of justice, which here is their capacity to honor fair terms of social cooperation, is represented by the various restrictions to which the parties are subject in the original position and by the conditions imposed on their agreement” (PL, 305).

88 TJ, 118.

89 TJ, 119.
There is nothing mysterious about this filtering process: it is simply meant to capture in a vivid way the considerations relevant and irrelevant to the choice of principles of justice. Finally, principles of justice are not selected from scratch, so to speak, but from “a short list of alternatives given by the tradition of moral and political philosophy,” the main alternatives in Rawls’s eyes being perfectionism, utilitarianism, and political liberalism.

§2.2.4 The distinction between a “political conception” and a “comprehensive doctrine”

The mature Rawlsian conception of the well-ordered polity—or political liberalism for short—is informed and regulated throughout by the distinction between political conceptions and comprehensive doctrines. This distinction is fundamental to political liberalism not only insofar as it informs Rawls’s analysis of the moral life of citizens, but also insofar as it profoundly shapes Rawls’s understanding of the horizon of possibilities and appropriate tasks of political philosophy itself. In short, the political vs. comprehensive distinction is not just a link in the argument of political liberalism: it is constitutive of the spirit animating Rawls’s entire project. Absent the political vs. comprehensive distinction, Rawls’s conception of political philosophy (and not just his own instantiation of it) would be unintelligible. By this, I do not mean that Rawls rejects the possibility of doing political philosophy in the “comprehensive” mode, so to speak, but that insofar as political philosophy offers itself as a vindication or critique of the basic structure of society, it ought to be “political, not metaphysical.” Political philosophy,

90 PL, 305.
insofar as it serves the public function of specifying and critiquing the principles of justice of our society, must abstract from deep philosophical disputes and make do, to the extent practicable, with the shared intuitions or considered moral beliefs of our society. Here is how Rawls puts it:

The aim of political philosophy, when it presents itself in the public culture of a democratic society, is to articulate and to make explicit those shared notions and principles thought to be already latent in common sense; or, as is often the case, if common sense is hesitant and uncertain, and doesn’t know what to think, to propose to it certain conceptions and principles congenial to its most essential convictions and historical traditions … I should emphasize that what I have called the “real task” of justifying a conception of justice is not primarily an epistemological problem. The search for reasonable grounds for reaching agreement rooted in our conception of ourselves and in our relation to society replaces the search for moral truth interpreted as fixed by a prior and independent order of objects and relations, whether natural or divine, an order apart and distinct from how we conceive of ourselves. The task is to articulate a public conception of justice that all can live with who regard their person and their relation to society in a certain way.91

The basic idea behind the political vs. metaphysical distinction is that we can distinguish on one hand elements of our worldview that constitute, or presuppose, philosophical or religious principles governing the moral life in general (not just political life), which elements taken together are considered either fully or partially “comprehensive”; and on the other hand, elements of our worldview that are not tied to any “comprehensive” picture or ideal of moral life, elements that speak more or less exclusively to our political relationships, which elements taken together are considered “political” rather than comprehensive. For Rawls, the fundamental characteristic

---

distinguishing comprehensive from political conceptions is their scope, that is, the domain of life for which they are intended, or in which they are considered relevant. On Rawls’s account, the content of a moral view or conception must differ substantially in virtue of its scope. Thus, the more general in scope a moral conception is, i.e. the more extensive the spheres of life to which it is meant to apply, the deeper and more philosophically complex it is likely to be. Stated simply, generality in scope requires comprehensiveness in content. In any case, the general assumption Rawls adopts is that as a theory aims to speak to wider domains of life (e.g. religious, metaphysical, psychological, etc.), it will undergo substantial changes; and that there are good reasons for the political theorist to view as especially significant the shift in content between theories that apply exclusively to the political domain and theories that apply to non-political domains as well. The former are called “political conceptions” while the latter are called “comprehensive doctrines.” Since this distinction is so central to political liberalism, it is worth presenting Rawls’s own explanation verbatim:

…the distinction between a political conception of justice and other moral conceptions is a matter of scope: that is, the range of subjects to which a conception applies and the content a wider range requires. A moral conception is general if it applies to a wide range of subjects, and in the limit to all subjects universally. It is comprehensive when it includes conceptions of what is of value in human life, and ideals of personal character, as well as ideals of friendship and of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole. A conception is fully comprehensive if it covers all recognized values and virtues within one rather precisely articulated system; whereas a conception is only partially

---

92 To avoid terminological confusion, I should point out that while Rawls tends to reserve the term “comprehensive” to refer to the content of conceptions and “generality” to refer to their scope, we could equally well use the term “comprehensive” to refer to scope and content.
comprehensive when it comprises a number of, but by no means all, nonpolitical values and virtues and is rather loosely articulated.\(^{93}\)

Rawls’s attempt to set political conceptions apart from comprehensive doctrines has come under fire from a variety of different directions. For example, he has been accused of attempting the impossible task of “de-politicizing” political theory by rendering its claims philosophically and religiously innocuous.\(^{94}\) I do not propose to explore these criticisms here. Instead, I propose to expose some of the basic assumptions Rawls has to make in order to sustain the political/comprehensive distinction, and then to explain briefly what theoretical and practical function this distinction serves in political liberalism.

First, the distinction between political and comprehensive conceptions presupposes that it is philosophically and psychologically tenable to develop moral ideals or principles that are tailored exclusively to one sphere of activity, and are considered intelligible and operational independent of wider or deeper conceptions of the moral life that apply to more extensive domains. In other words, we must, if Rawls’s account is right, be capable of reasoning practically about one domain, in this case politics, without allowing principles applied to other domains, e.g. religion or ethical life more generally, to intrude upon or “infect” our reasoning in the limited domain. This is not to say that we may not privately uphold the political conception for reasons drawn from our

\(^{93}\) PL, 13.

\(^{94}\) Although Connolly’s *The Terms of Political Discourse* was written before political liberalism was developed, it already contains an implicit critique of the whole enterprise of seeking to avoid philosophical controversy in political theory. For a more explicit critique of Rawls’s attempt to do political theory without the baggage of metaphysics, see Richard Bellamy and Martin Hollis, ”Liberal Justice: Political and Metaphysical,” *The Philosophical Quarterly* 45, no. 178 (1995); J. Judd Owen, *Religion and the Demise of Liberal Rationalism: The Foundational Crisis of the Separation of Church and State* (Chicago: The University of Chicago Press, 2001), esp. chap. 5, “Religion and Rawls’s Freestanding Liberalism.”
comprehensive doctrine, but it is to say that we must be able to publicly affirm and debate the political conception and its implications, in good conscience, without relying time and again on our comprehensively-derived reasons. In other words, our practical reasoning in the domain of politics must be capable of yielding valid and effective arguments for this or that course of action, independently of the more comprehensive values that inform the rest of our life.\textsuperscript{95}

It was not until Rawls’s essay, “Justice as Fairness: Political not Metaphysical,”\textsuperscript{96} that the political vs. comprehensive distinction was explicitly developed. Whether it is ultimately tenable or not, it is clear that it serves a crucial function in the shift from Theory to Political Liberalism. Rawls admitted in the introduction to the paperback edition of Political Liberalism that Theory of Justice had effectively embedded justice as fairness in a Kantian moral view, and thus rendered it too philosophically partisan to be amenable to an “overlapping consensus” of citizens. Rawls believed that in Theory, he had underestimated the depth of moral and religious diversity afflicting liberal democratic cultures (the “fact of reasonable pluralism”), and that justice as fairness had to be framed differently if it were to have any hope of winning the allegiance of citizens across religious and moral divides. Much of the substance of justice as fairness, including the original position, could be preserved, but it could no longer be based on a general moral view that applies across the gamut of human endeavors. Instead, it must be reinterpreted as a moral view with a limited scope or domain (politics), with no ambition

\footnote{\textsuperscript{95} In fact, Rawls goes so far as to say that the political conception should be presented as agnostic about its own truth or falsehood: it should be presented as reasonable rather than true. For a critique of Rawls’s “epistemic abstinence,” see Raz, "Facing Diversity."}

\footnote{\textsuperscript{96} Rawls, "Justice as Fairness: Political Not Metaphysical."}
to make sense of the whole of morality. The newly conceived constraints on scope required some modifications to the content of justice as fairness, principally to its account of the stability of the well-ordered society, but most elements of the theory, according to Rawls, are consistent with Political Liberalism. The main shift inaugurated by political liberalism is not in the internal structure of justice as fairness, but how it is viewed and grounded. Justice as fairness is no longer viewed as a partially comprehensive doctrine, but as a “political conception,” which speaks exclusively to our activities as citizens or members of a liberal democracy, not to our pursuit of the good more generally. Furthermore, justice as fairness is not the only political conception of justice, but one proposed by Rawls as “the most reasonable” of a “family of political conceptions.”

§2.2.5 The distinction between public and nonpublic reason

Not only Rawls’s “freestanding” conception of political philosophy, but also his view of the conditions and norms governing public political discourse, builds off, and presupposes, the distinction between political and comprehensive doctrines. The term “public reason” is used by Rawls to refer to a certain domain of reasoning that is of special moral significance for the political philosopher, namely, that domain of reasoning

---

97 See PL, introduction, xviii: “the serious problem I have in mind concerns the unrealistic idea of a well-ordered society as it appears in Theory. An essential feature of a well-ordered society associated with justice as fairness is that all its citizens endorse this conception on the basis of what I now call a comprehensive doctrine. They accept, as rooted in this doctrine, its two principles of justice...Political liberalism [on the other hand] assumes that, for political purposes, a plurality of reasonable yet incompatible comprehensive doctrines is the normal result of the exercise of human reason within the framework of the free institutions of a constitutional democratic regime.”


99 As Rawls puts it, “the content of public reason is given by a family of political conceptions of justice, and not by a single one.” (Rawls, "The Idea of Public Reason Revisited," 773).
pertaining to law and public institutions, insofar as they coercively structure people’s lives. Domains of reason may be distinguished according to the goal of the reasoning and its subject-matter: for example, we may speak of “scientific” reason, “religious” reason, “philosophical” reason, “juridical” reason and so on, as we move between reasoning with scientific, religious, philosophical, and juridical subjects and goals in mind. Different spheres of reason are governed by different rules, and what counts as competent and responsible reasoning in one sphere is not identical to what counts as competent and responsible reasoning in another. For example, a judge in his juridical deliberations, must discount evidence that was seized illegally, even if de facto, it proves beyond a reasonable doubt the guilt of a violent criminal; whereas as a private citizen, it would be perfectly legitimate for the judge to take precautions or be wary (in light of the discounted evidence) in his dealings with the defendant. ¹⁰⁰

What distinguishes “public” reason from other spheres of reason, as stipulated by Rawls, is not (at least not primarily) that it is open to view or accessible to citizens, but that it concerns the coercive power that citizens can, either directly or through their political representatives, exercise over one another. As we have already seen, the Rawlsian liberal concept of citizens as free and equal, combined with the fact of reasonable pluralism, makes coercion, on its face, morally suspect. It is on this basis that Rawls pays special attention to those exercises of reason that conclude in acts of public coercion. As Rawls puts it, public reason is “the reason of equal citizens who, as a collective body, exercise final political and coercive power over one another in enacting

¹⁰⁰ Although the second element of this contrast is not given by Rawls, the example of the discounting of illegally seized evidence by a judge is. See PL, 218.
laws and in amending their constitution."\textsuperscript{101} The reason of associations (say, the Boy Scouts, or the Catholic Church), even though it may formulate sanctions including expulsion for noncompliance with the association’s rules, is considered nonpublic to the extent that it is not strictly coercive, once we assume that membership in the association is optional rather than obligatory.\textsuperscript{102}

As I have already suggested, the domain of public reason is specified on normative grounds, viz., the coercive dimension of government and law. But Rawls clearly wants to go further than just specifying which areas of reasoning have these special coercive consequences; he also wants to investigate which moral principles ought to guide this sort of reasoning. For convenience, we might distinguish between the domain of public reason—those questions and claims that fall within its scope—and the full-blooded ideal of public reason, which specifies the norms or rules which govern its exercise.

The ideal of public reason requires a somewhat elaborate theoretical apparatus. Essentially, it proposes some fairly simple rules for the conduct of public reason, but rules that are only intelligible in light of certain important philosophical assumptions. Some of the rules of public reason are rules that pertain to reason more generally, e.g. the law of non-contradiction and other rules of logical inference. However, what interests us here are those rules that are distinctive of the ideal of public reason. Probably the central requirement of Rawls’s ideal of public reason can be stated in two steps: first, and most

\textsuperscript{101} PL, 214.

\textsuperscript{102} Of course, it might be argued that the exit costs of some associations (e.g. social stigma) are prohibitively high and thus that membership is in effect compulsory. However, this is clearly a matter of degrees, whereas the case for treating membership in the State as practically compulsory seems more clear-cut, given the consistently high costs associated with emigration.
fundamental, is the requirement to offer reasons (or arguments) that other citizens could reasonably accept.\textsuperscript{103} Second, to be inferred from the first, is the requirement to offer, or be prepared to offer, reasons (or arguments) that are based on a political conception, rather than on a comprehensive doctrine.\textsuperscript{104} This inference is not one of immediate logical entailment, of course: on the contrary, it requires the acceptance of the distinction between political and comprehensive conceptions, and it presupposes that political conceptions meet at least two conditions: (a) they are a source of reasons (or arguments) that citizens could reasonably accept, in a way that sets them apart from comprehensive doctrines; and (b) they provide a set of reasons or arguments adequate to the task of debating and justifying the relevant laws or policies.\textsuperscript{105} Furthermore, the major premise of the argument, the proposition that citizens are required to offer reasons that other citizens “could reasonably accept,” is only intelligible in light of Rawls’s (political) conception of the person, as we have already seen (§2.2.2 above).

\textsuperscript{103} This is the essence of the principle of reciprocity on my interpretation. See §3.1.2 and §3.2 below for a discussion of reciprocity and its philosophical foundations.

\textsuperscript{104} This requirement, which Rawls calls the “duty of civility” has been qualified by Rawls in response to criticisms, but in its essentials it remains intact. According to the so-called “proviso,” for example, nobody is morally precluded from advocating laws based on nonpublic reasons, but they ought to find public reasons “in due course.” It is not clear to me how the so-called “proviso” of “in due course” significantly softens the duty of civility. After all, surely a citizen cannot advance nonpublic reasons unless he or she at least has grounds for believing, in good faith, that comparatively persuasive public reasons will be forthcoming sooner rather than later? I would contend that to require that citizens put so much faith in public reason and its capacity to support their political beliefs does not significantly mitigate the demandingness of the duty of civility.

\textsuperscript{105} As Rawls puts it, “the substantive content and the guidelines of inquiry of a political conception” should be “complete,” meaning that “the values specified by that conception can be suitably balanced or combined, or otherwise united, as the case may be, so that those values alone give a reasonable public answer to all, or to nearly all, questions involving the constitutional essentials and basic questions of justice” (PL, 225).
§2.2.6 The Idea of a Well-Ordered Society

In §2.1, “Ideals of citizenship,” I suggested some reasons for believing that ideals of citizenship and their practical demands are partially shaped by contingent factors such as socio-cultural conditions and morally relevant features of the case at hand. Rawls’s ideal of citizenship is embedded within, or presupposes, a certain conception of social order, as we shall see when we explain the Rawlsian ideal of citizenship in the next section. Rawls calls this conception “the idea of a well-ordered society.” The basis of the ideal is, of course, Rawls’s political conception of justice, constituted by the principles of justice and the ideal of public reason. However, the idea of the well-ordered society specifies more explicitly (than the political conception of justice) the paradigmatic instance in which Rawls’s political conception of justice holds unproblematically:

To say that a society is well-ordered conveys three things: first (and implied by the idea of a publicly recognized conception of justice), it is a society in which everyone accepts, and knows that everyone else accepts, the very same principles of justice; and second (implied by the idea of the effective regulation of such a conception), its basic structure—that is, its main political and social institutions and how they fit together as one system of cooperation—is publicly known, or with good reason believed, to satisfy these principles. And third, its citizens have a normally effective sense of justice, and so they generally comply with society’s basic institutions, which they regard as just. In such a society the publicly recognized conception of justice establishes a shared point of view from which citizens’ claims on society can be adjudicated.106

If a society is well-ordered, or at least approximately well-ordered in the relevant sense, then Rawls’s ideal of justice and public reason, and the associated ideal of citizenship, presumably hold sway—indeed, if they do not hold sway under these optimal conditions, when could they hold sway? However, if a society falls short of being well-

106 PL, 35.
ordered in the relevant sense, or only partially approximates the well-ordered society, it is much less clear which parts, and to what extent, Rawls’s conception of justice should be followed. For example, in the real world many citizens do not in fact regard their society’s basic institutions as just. Does this mean that citizens are no longer bound by the duty of civility? Or does it mean that they may disregard the second principle of justice? Political liberalism stipulates that society is “well-ordered,” and Rawls unfortunately devotes very little discussion to the complications of applying a highly idealized theory to a messy and only imperfectly just social and political order.

If the idea of the well-ordered society presupposed by political liberalism is, by Rawls’s own admission, “highly idealized,” then one might wonder what possible real-world check there can be on the theory. If it succeeds by definition or stipulation, then this is a very cheap success indeed. Here, it is helpful to introduce Rawls’s distinction between a “utopian” ideal and a “realistically utopian” ideal. The litmus test for the success of Rawls’s theory is not only whether it cogently matches and orders our “considered convictions” about justice, but whether it presents an ideal which we could imagine our society, under reasonably (but not unrealistically) favorable conditions, approximating. Is a well-ordered society regulated by a political conception of justice within the range of possibilities we could realistically contemplate our society approximating in the foreseeable future? If so, then the ideal of political liberalism is “realistically utopian” and we can somehow work towards it; if not, it is utopian pure and

107 For the three conditions characterizing a “well-ordered” society, see PL, 35 (quoted above).

108 PL, 35.
simple, and we had better not waste any further energy on it (however theoretically illuminating it might be).  

Helpful as the distinction between utopian and realistically utopian theories is, we are not quite out of the forest yet. For even supposing that political liberalism is “realistically utopian,” there remains the thorny question—which Rawls all but ignores—how rigorously ought we allow it to regulate our conduct as we work towards it in a society which is only imperfectly well-ordered? If a significant number of my fellow citizens are not prepared to observe the duty of reciprocity in their dealings with me, e.g. are not prepared to honor their commitments, then it may be that I am absolved from the duty myself. On this particular point, Rawls is relatively clear. However, other points are much less clear, e.g. if many of my fellow citizens routinely violate the duty of civility in public discourse, am I still bound to observe it? Or if the duty of civility is contingent upon reciprocity, what is the threshold compliance level at which the duty kicks in? I am not suggesting that the failure to address these questions is necessarily fatal to political liberalism, nor that any clear rules could settle such questions in advance. However, they are questions that probably deserve more discussion than they receive from exponents of political liberalism, bearing as they do on the normative status or relevance of political liberalism in the world we actually inhabit. For the most part, my

---

109 For the notion of a “realistically utopian” theory, see Rawls, *Justice as Fairness: A Restatement*, 4: “We view political philosophy as realistically utopian: that is, as probing the limits of practicable political possibility. Our hope for the future of our society rests on the belief that the social world allows at least a decent political order, so that a reasonably just, though not perfect, democratic regime is possible. So we ask: What would a just democratic society be like under reasonably favorable but still possible historical conditions, conditions allowed by the laws and tendencies of the social world?”

110 The requirement of reciprocity is expressed as a condition attaching to cooperation: “fair terms of cooperation... are terms that each participant may reasonably accept, provided that everyone [or practically everyone] else likewise accepts them” (PL, 16, emphasis added).
treatment of political liberalism will concentrate on its merits as a liberal democratic ideal, holding certain fundamental social conditions constant, rather than the logic of its implementation in an imperfectly well-ordered society.

§2.3 The Rawlsian Ideal of Citizenship

So far, I have clarified some aspects of the notion of an ideal of citizenship, narrowing the meaning of the term in accordance with Rawls’s theoretical and practical purposes; and I have briefly explicated the fundamental premises of Rawls’s conception of the well-ordered polity. With these aspects of Rawls’s theoretical apparatus in view, we are now in a much better position to appreciate the full content and significance of Rawls’s ideal of citizenship. But before proceeding, let us briefly recollect our working definition of an ideal; and the fundamental premises of political liberalism according to my interpretation. Recall that an ideal of citizenship is, roughly, (a) a pattern of morally desirable action and dispositions pertaining to the special role of citizen, (b) with superogatory as well as deontological connotations, (c) conditioned by both contingent factors such as social and political structures and conventions, and invariant factors such as human nature. And political liberalism is motivated primarily by the goal of providing an acceptable basis for citizens to live peacefully together in a pluralistic society, which means a basis that is (i) principled, not merely pragmatic; (ii) derived from reasoned, not coerced agreement; and (iii) sufficiently supported by citizens to be stable, not precarious. Finally, political liberalism is constituted by six fundamental premises (i) the “fact of reasonable pluralism”; (ii) the concept of the citizen as free and equal; (iii) the concept of society as a fair and stable system of social cooperation enduring over time;
(iv) the distinction between a “political conception” and a “comprehensive doctrine”; (v) the distinction between public and nonpublic reason; and (vi) the idea of a well-ordered society.

§2.3.1 The Place of the Ideal of Citizenship in Political Liberalism

Now, the Rawlsian ideal of citizenship is not easily isolated from the rest of Rawls’s theory. This is because what it takes to be a good citizen is not simply a logical consequence of political liberalism, but, at least in part, present in the bedrock assumptions that drive the entire theory. In this sense, Rawls’s ideal of citizenship can be thought of as embedded within the foundations of his theory rather than a logical outgrowth or something peripheral. For example, one of the most important premises of political liberalism is the proposition that citizens are free and equal, in the sense that they are equally entitled to enjoy the conditions necessary for the adequate development of their two moral powers (the capacity for a conception of the good and the capacity for a sense of justice). This entitlement, and the corresponding duties imputed to society collectively, already begin to define norms of citizenship (which together constitute, or partially constitute, an ideal of citizenship), e.g., respect the entitlement of others to enjoy the conditions necessary for the adequate development of the two moral powers. Thus, elements of the Rawlsian ideal of citizenship are already present at the deepest level of his theory.

Rawls’s ideal of citizenship, i.e. his prescribed pattern of excellences pertaining to the role of citizen, is informed by one key requirement of political liberalism, which I will label the requirement of respect for individual autonomy. We are obligated to respect
the moral autonomy of our fellows\textsuperscript{111}: Each of us is entitled to the social conditions necessary for the adequate development and exercise of our two moral powers, viz. the capacity for a conception of the good, and the capacity for a sense of justice. When these two moral powers are realized in one’s life, one is living a “fully” autonomous life. Autonomy is not a description of the good life but a description of the basic conditions under which anyone can pursue a meaningful and dignified human existence. It is important to note here that Rawls does not mean by autonomy anything as exalted as, say, the autonomy of Kant’s noumenal self: rather, Rawls has in mind the ability of citizens to politically determine their life together on mutually acceptable terms. This is what Rawls calls “political” as opposed to “ethical” autonomy.\textsuperscript{112}

On the other hand, it is not the case that Rawls assumes a formal ideal of citizenship \textit{a priori} or independent of empirical conditions. Thus, when Rawls imagines the well-ordered society, he must also ascertain which character traits and attitudes are conducive to political stability and collaboration under favorable conditions, and which are not. This is only something we can know based on generalizations or extrapolations

\begin{quote}
\textsuperscript{111} Here, I should note that I am not using the qualification “moral” in the special sense Rawls uses it to distinguish “moral” or “ethical” from “political” autonomy, but in the more general sense of, “pertaining to the moral life or to human action.”
\end{quote}

\begin{quote}
\textsuperscript{112} The meaning of this distinction is explained in PL, 77-78: “Here I stress that full autonomy is achieved by citizens: it is a political and not an ethical value. By that I mean that it is realized in public life by affirming the political principles of justice and enjoying the protections of the basic rights and liberties; it is also realized by participating in society’s public affairs and sharing in its collective self-determination over time. This full autonomy of political life must be distinguished from the ethical values of autonomy and individuality, which may apply to the whole of life, both social and individual, as expressed by the comprehensive liberalisms of Kant and Mill. Justice as fairness emphasizes the contrast: it affirms political autonomy for all but leaves the weight of ethical autonomy to be decided by citizens severally in light of their comprehensive doctrines.” Of course, whether or not it is coherent to isolate autonomy in the political sphere from autonomy in the ethical sphere as sharply as Rawls does, is an interesting question we need not consider here, except to say that our assessment of this debate will turn to a large extent on how persuasive we find the distinction between political and comprehensive doctrines and its function in political liberalism.
\end{quote}
from empirical findings and our own experience. In this respect, Rawls repudiates Kant’s attempt to isolate the principles of practical reason from the “phenomenal” or empirical world.\textsuperscript{113} It would appear, then, that Rawls’s ideal of citizenship is partly constituted directly by the fundamental normative commitments of political liberalism, and partly derived from its theoretical commitments via the empirical conditions necessary for those commitments to be made good.

\textbf{§2.3.2. Reasonableness and the Cooperative Virtues of Citizenship}

Assuming, as I have argued, that good citizens are expected to respect the autonomy of their fellow citizens, there are certain derivative rules and stable dispositions or virtues which might be thought of as ultimately ancillary to, or reducible to, the requirement of respect for the individual.\textsuperscript{114} For example, Rawls expects citizens to be amenable to accepting fair terms of cooperation that could be accepted by others, and this disposition is captured by the notion of “reasonableness.” Reasonableness is the disposition to meet others half-way, and to make a reasonable effort to live on fair terms acceptable to all. This is the basic precondition for a society where each individual is capable of, or at least has the opportunity, to live according to his or her conception of the good and sense of justice, at least subject to the reasonable limitations of life in society.

\textsuperscript{113} According to Rawls, it is obviously impossible to develop a substantive theory of justice founded solely on truths of logic and definition: “The analysis of moral concepts and the a priori, however traditionally understood, is too slender a basis. Moral theory must be free to use contingent assumptions and general facts as it pleases. There is no other way to give an account of our considered judgments in reflective equilibrium” (TJ, 44).

\textsuperscript{114} This is clear once we assume that the mark of respect for another, at least within the political domain, is the disposition to treat him or her in accordance with justice. As Rawls puts it in \textit{Theory}, “once the principles of right and justice are on hand, they may be used to define the moral virtues just as in any other theory. The virtues are sentiments, that is, related families of dispositions and propensities regulated by a higher-order desire, in this case a desire to act from the corresponding moral principles” (TJ, 167).
Although Rawls does not propose reasonableness as a “master” virtue, I believe it could be viewed as the “master” virtue of cooperativeness, or the general disposition to cooperate with others, at least where others are willing to reciprocate.\textsuperscript{115} By “master” virtue, I mean something like a cardinal virtue, that is, a virtue which is composed of, or entails, a series of dispositions that can be described one by one, or left implicit within the master virtue. Thus, another example of a master virtue would be fortitude, or the disposition to bear difficulties or trials with courage. This clearly requires a series of dispositions, including patience, temperance or self-control, and prudence. Of course, some of these sub-virtues may well be viewed as master virtues themselves, e.g. temperance, and they may have their own characteristic sub-virtues. Thus, their status as master virtue or sub-virtue will sometimes depend on how we are viewing them.

Reasonableness is not only the master virtue of cooperativeness for Rawls, but the only master virtue in Rawls’s ideal of citizenship. Indeed, Rawls appears to argue that no virtue falling outside the domain of reasonableness is admissible within the political ideal of citizenship, since this would jeopardize the political, as opposed to comprehensive, status of the ideal:

\textsuperscript{115}Reasonable persons, “when they believe that institutions or social practices are just, or fair……are ready and willing to do their part in those arrangements provided they have reasonable assurance that others will also do their part” (PL, 86). The more complex and—arguably—the more problematic description of the reasonable person runs like this: “Citizens are reasonable when, viewing one another as free and equal in a system of social cooperation over generations, they are prepared to offer one another fair terms of social cooperation (defined by principles and ideals) and they agree to act on those terms, even at the cost of their own interests in particular situations, provided that others also accept those terms. For these terms to be fair terms, citizens offering them must reasonably think that those citizens to whom such terms are offered might also reasonably accept them … And they must be able to do this as free and equal, and not as dominated or manipulated, or under the pressure of an inferior political or social position” (PL, introduction to the paperback edition, xliv, emphasis added).
Even though political liberalism seeks common ground and is neutral in aim, it is important to emphasize that it may still affirm the superiority of certain forms of moral character and encourage certain moral virtues. Thus, justice as fairness includes an account of certain political virtues—the virtues of fair social cooperation such as the virtues of civility and tolerance, of reasonableness and the sense of fairness (IV:5-7). The crucial point is that admitting these virtues into a political conception does not lead to the perfectionist state of a comprehensive doctrine.116

One obstacle to my interpretation of reasonableness as the “master virtue” of Rawls’s ideal of citizenship is that Rawls occasionally speaks of reasonableness as if it were just one more virtue on a par with the other cooperative virtues. This is what we find in the passage I just quoted, where reasonableness is listed, without any special distinction, alongside other cooperative virtues, namely civility, tolerance, and a sense of fairness. Yet if we refer back to Rawls’s own account of the virtue of reasonableness, it becomes clear that it is a virtue on a different level to other cooperative virtues, since its content effectively encompasses, albeit implicitly, the content of all the other cooperative virtues listed by Rawls. “Citizens are reasonable,” says Rawls, “when, viewing one another as free and equal in a system of social cooperation over generations, they are prepared to offer one another fair terms of social cooperation (defined by principles and ideals) and they agree to act on those terms, even at the cost of their own interests in particular situations, provided that others also accept those terms.”117 In defending his account of civic or political virtue against perfectionist misgivings, Rawls insists that “the ideals connected with the political virtues are tied to the principles of political justice and to the forms of judgment and conduct essential to sustain fair social cooperation over

---

116 PL, 194, emphasis added.
117 PL, xlv.
And what is reasonableness, if not the disposition to engage in “fair social cooperation over time”? It seems to me that the most illuminating way of reading Rawls’s catalogue of the virtues is to say that reasonableness involves the overall disposition to engage in fair cooperation with others viewed as free and equal, whereas the individual virtues, such as civility and tolerance, refer to more specific aspects or components of the general disposition towards fair cooperation. For example, the disposition to trust those who have demonstrated their moral credentials is an aspect of the disposition to engage in fair cooperation with others (reasonableness), while the disposition to treat others with civility or courtesy, or to show respect for others’ equality and freedom in one’s informal dealings with them, is likewise an essential aspect of the disposition to engage in fair cooperation with others (viewed as free and equal citizens).

Few of the cooperative virtues are developed at length by Rawls, but many of them are at least mentioned: The “morality of association,” the third stage in Rawls’s account of moral development given in Theory, is characterized by the “cooperative virtues,” namely “those of justice and fairness, fidelity and trust, integrity and impartiality.” The associated vices are “graspingness and unfairness, dishonesty and deceit, prejudice and bias.” Cooperative virtues mentioned elsewhere by Rawls are “the

---

118 PL, 194.

119 The three stages of moral development in Rawls’s account are based on Lawrence Kohlberg’s account of the moral development of persons from childhood through adolescence (cited in TJ, fn. 6, p. 403). The three stages are (i) the morality of authority; (ii) the morality of association, and (iii) the morality of principles (TJ, 405-419). The details need not concern us here. Suffice it to say that “[w]hereas the child’s morality of authority consists largely of a collection of precepts, the content of the morality of association is given by the moral standards appropriate to the individual’s role in the various association to which he belongs. These standards include the common sense rules of morality along with the adjustments required to fit them to a person’s particular position; and they are impressed upon him by the approval and disapproval of those in authority, or by the other members of the group” (TJ, 409). I do not believe Rawls’s ideal of citizenship, or the virtues associated with it, presupposes or depends upon this particular account of moral development, though some such account would seem necessary to explain the sociological origins of the virtues and the motivation to adhere to the ideal of citizenship.
virtues of civility and tolerance.”⁹¹²⁰ The basis for calling these the virtues of cooperation is that they may be regarded as “forms of judgment and conduct essential to sustain fair social cooperation over time.”⁹¹²¹ In other words, to claim to be reasonable or cooperative in general but to lack one of the cooperative virtues would be incoherent.

One area of human activity that is especially relevant to cooperation is, of course, political argumentation and debate, in particular, debate that issues in legal and policy determinations. In this arena of human activity, Rawls speaks of an “ideal of public reason,”⁹¹²² or a view about how we ought to conduct ourselves as we reason in public about coercive laws and policies. Given the value of individual autonomy which is to be protected and supported by the polity at large, we ought to reason about policies and laws in a way that is compatible with respect for the full autonomy of our fellow citizens. This means that we need to advance reasons in this arena that are, if not actually accepted by others, at least “accessible” or “acceptable” to them. For Rawls, this means that the reasons we advance ought to be reasons others “could reasonably accept.” This is a condition which is not unpacked very far by Rawls, but for now, I take it to mean that we must imagine ourselves in the shoes of another and see whether he could, upon reflection, at least potentially and without excessively straining his current epistemic commitments, come to accept the reason in question as a valid and justifying reason for the action being proposed.⁹¹²³ Notice that this does not mean that our interlocutor would agree with the

---

⁹¹²⁰ PL, 194.

⁹¹²¹ PL, 194.

⁹¹²² For a good statement of this ideal, see PL, introduction, If.

⁹¹²³ Of course, reasons do not come in isolation from wider theories and assumptions. Political reasons, for example, tend to be embedded within a political conception of justice. Thus, insofar as a reason is
reason, but we could at least imagine him coming to be persuaded without converting to a different comprehensive doctrine. Indeed, in many cases, the addressee himself might reject as untenable a reason offered, yet still view it as epistemically accessible and politically relevant in the case at hand, the sort of reason he could imagine himself being persuaded by without converting to a fundamentally different weltanshaaung. Anything short of this would be morally unacceptable, since it would violate a person’s right to pursue a conception of the good of his choice within the bounds of public order (one dimension of his autonomy). Hence, Rawls asserts that the “duty of civility” or the “ideal of public reason” requires that “citizens are to conduct their fundamental discussions within the framework of what each regards as a political conception of justice based on values that the others can reasonably be expected to endorse and each is, in good faith, prepared to defend that conception so understood.”

Unfortunately, aside from a few cases such as the sense of fairness and trust, and some narrowly specified cases of virtue such as the “duty of civility,” Rawls does not flesh out for us any rich account of the cooperative virtues. For example, there is only an application of civility to the special case of public reason; there is some discussion of cases of tolerance but little positive analysis of the virtue as such; and fidelity and integrity are hardly discussed at all. The theoretical and practical inadequacy of a partial

---

124 PL, 226. Some have found Rawls’s reciprocity criterion unduly vague and ambiguous, and particularly susceptible to ideological manipulation. See, for example, Robert P. George and Christopher Wolfe, “Natural Law and Public Reason,” in Natural Law, Liberalism, and Morality, ed. George and Wolfe, 51-74. I believe Rawls could have said a lot more to clarify what he had in mind by reciprocity, and I attempt to make the notion more concrete in §3.1, relying in part on Eberle’s discussion of public justification in Religious Convictions in Liberal Politics, esp. ch. 3, “Justificatory Liberalism.”

125 See TJ, “Toleration of the Intolerant,”190-194; PL, xxvi f., 194 f.
account of civic virtue such as this one will only become fully apparent in light of my broader treatment of civic virtue in chapter 6. Rather than guessing what Rawls would have said had he further developed his account of the cooperative virtues, I only mean to point out the incompleteness while emphasizing that the account as a whole derives its integrity from the master virtue of reasonableness. But wherein lies the value of this “master” virtue?

There is something about the central role of autonomy in Rawls’s conception of justice and political right that tends to foreground the cooperative virtues (under what I would term the umbrella virtue of “reasonableness”), or at least a certain version of them, and render other candidates less eligible for the virtues of citizenship. This will become clearer when we subject Rawls’s conception of civic virtue to more critical analysis in §5.2 and §5.3. For now, suffice it to say that the central requirement that we respect the autonomy of our fellows, understood as the capacity for the exercise of the two moral powers and the conditions for their development, and the associated primacy of political over comprehensive values, naturally requires and validates the cooperative virtues, whereas it brings into question the political, or public normative status of virtues more loosely tied to voluntary cooperation, such as humility, modesty, and courage.126

***

126 Notice that these virtues, unlike, say, the virtue of “reasonableness,” are more readily associated with “comprehensive” conceptions of human flourishing such as (some versions of) civic republicanism or neo-Aristotelian virtue ethics, that extend far beyond the strictly political sphere.
In this chapter, I have laid some of the essential conceptual groundwork for both the critical and constructive phases of the dissertation. First, I have attempted to articulate a plausible interpretation of the notion of an *ideal* of citizenship, that is both informative and practically neutral between classical and modern orientations. I will assume this conception of an *ideal* of citizenship when it comes time to outline an alternative approach to citizenship in chapter 6. Secondly, I have exposed the basic grounds and content of Rawls’s ideal of citizenship, providing a foundation both for my critique of the duty of civility (chapter 3), and for an evaluative comparison between contractualist ideals and a richer aretaic ideal of citizenship (chapters 5 and 6). In short, this chapter aims to provide a solid foundation both for critiquing and improving on Rawls’s attempt to solve the problem of citizenship in contemporary democratic cultures.
CHAPTER 3: A CRITIQUE OF PUBLIC REASON FROM INTEGRITY AND RECIPROcity

Now that I have explicated the most fundamental elements of the Rawlsian vision of political order and explored their inter-connections, remaining as close as possible to the spirit and intention of Rawls himself, I propose to subject the Rawlsian ideal of citizenship to closer and more critical scrutiny. Rawls has especially come under criticism for the so-called “duty of civility,” which imposes a moral duty upon citizens to advance publicly accessible arguments for positions bearing on constitutional essentials or matters of basic justice. But these criticisms tend to come either from an alternative strand of liberalism\textsuperscript{127}, or from an explicitly religious perspective.\textsuperscript{128} I propose to critique Rawls’s ideal of public reason both from an external, integralist, perspective (a perspective that affirms the value of ethical integrity), and from a strictly Rawlsian perspective. That is, in addition to considering the implications of public reason for ethical integrity, I will evaluate it firmly within the terms of political liberalism itself,

\textsuperscript{127} See, for example, Ronald Dworkin, "Foundations of Liberal Equality", in The Tanner Lectures on Human Values (Salt Lake City: University of Utah Press, 1990), 1-120; Christopher Eberle, Religious Convictions in Liberal Politics (Cambridge: Cambridge University Press, 2002); and Michael J. Perry, "Why Political Reliance on Religiously Grounded Morality is not Illegitimate in a Liberal Democracy", Wake Forest Law Review 36 (Summer 2001): 217-249.

rather than in the terms of religion, ethics more broadly speaking, or some alternative
strand of liberalism.

Before I proceed, it is only fair that I acknowledge a deep intellectual debt to John
Rawls. Heated disagreements and hard-hitting arguments are essential to a vibrant
intellectual community, but all too often obscure how much we share with, and even
learn from, our adversaries. As somebody who prizes personal freedom and a broad range
of political and civil liberties, I cannot help admiring the inclusive and non-coercive
spirit which surfaces time and again in Rawls’s thought. Rawls has done more than most
political philosophers of our time\textsuperscript{129} to draw attention to the need to square our political
commitments with the depth of moral disagreement that characterizes liberal democratic
societies. Furthermore, even if the particulars of Rawls’s explanation of disagreement are
controversial, most of us (including myself) accept his claim that frequently, moral and
religious disagreement is compatible with rationality and upright intentions, and as such,
cannot be automatically treated as the outcome of moral perversion or irrationality. That
there is such thing as reasonable disagreement in politics—even if not in the precise sense
meant by Rawls—and that we should make every reasonable effort to resolve it through
rational and fair procedures, is something I take as hardly in need of demonstration.

It is also worth mentioning at the outset of my critique that Rawls’s work has
done a great deal to highlight three fundamental problems confronting modern political
life: the problem of how best to show respect towards others in a pluralistic and free
society; the problem of how to ground the institutions of a free society in a way that is
stable or wins the allegiance of a preponderance of citizens; and the problem of
reconciling one’s public commitments as a citizen with one’s deepest ethical and

\textsuperscript{129} Alasdair MacIntyre, one of liberalism’s most trenchant critics, ironically shares this honor with Rawls.
religious convictions. Though I clearly have fundamental and far-reaching disagreements with Rawls’s peculiar articulation and response to these problems, I see myself as entering into a conversation with Rawls about the very problems he himself has spent so much time and energy thinking through. Thus, I hope it will be clear to the reader that taking Rawls’s thought as the backdrop for my own, even while vehemently disagreeing with it, is one of the highest compliments I can pay to another political philosopher.

This chapter will address two fundamental questions, the first grounded in concerns internal to political liberalism, and the second in concerns external to it: first, is the constraint political liberalism places upon civic discourse consistent with its own internal commitments? Second, what are the implications, positive or negative, of the Rawlsian ideal for citizens’ integrity? The thought motivating these questions is that Rawls’s efforts to impose moral (not legal) constraints on political discourse appears to lead him to a position that should be embarrassing to him given the explicitly tolerationist impulse of his justificatory project. The thesis which this chapter will defend is that political liberalism’s peculiar strictures on discourse130 and their deleterious implications for integrationist citizens131 are not only embarrassing in light of the commitments to freedom and toleration of difference that are characteristic of liberal theories in general, but in light of core commitments of Rawls’s own peculiar brand of liberalism.

---

130 I say peculiar strictures on discourse, to indicate that I am not arguing tout court against the practical necessity of a moral framework of expectations and norms to govern discourse in a liberal democracy. It is political liberalism’s peculiar strictures on discourse, not strictures on discourse as such, that are (or should be) embarrassing to any theorist hailing from the liberal tradition.

131 The distinction between integrationist and non-integrationist citizens is explained in §3.3.1. By integrationist citizens, I mean, roughly, citizens who seek to integrate the values pursued across the entire spectrum of activities in their life, be they family life, friendship, love, art, or politics.
I will begin by setting out in its essentials the content and rationale of Rawls’s so-called “duty of civility” which lies at the heart of his ideal of public reason (and, indeed, his ideal of citizenship) (§3.1). Secondly, I will argue from within political liberalism that the very premises allegedly underpinning the duty of civility—namely, the values of reciprocity and political autonomy, and the burdens of judgment—in fact rule it out. The conclusion of this argument is that this duty ought to be rejected by a Rawlsian, insofar as it violates the (political) autonomy of citizens (§3.2). Thirdly, I will argue that public reason unfairly or unreasonably burdens the integrity or conscientious pursuit of the good of reasonable citizens (§3.3). Fourthly, I argue that the duty of civility jeopardizes one of the central goals of political liberalism, namely political stability (§3.4). Finally, I argue that we have no reason to believe there are any interesting and “complete” political conceptions of justice that can meet the reciprocity requirement, at least as that requirement is interpreted by Rawls. Therefore, I suggest it is time to rethink not only the duty of civility, but the reciprocity requirement that underlies it (§3.5).

§3.1. The Duty of Civility

§3.1.1 Content of the Duty: Public Reason

Like any liberal, Rawls focuses on the fact of coercion as one of the morally salient features of a political system. Specifically, in a liberal democracy—at least one in which the institutional mechanisms of representation are in good order—coercion can be thought of as the collective power citizens wield over each other, through the
instrumentality of government and its agents. But the question immediately arises, assuming with Rawls that citizens are free and equal, how can the exercise of this collective power be justified to the citizens who feel its weight, whether directly in the form of penalties (whether applied or threatened) or indirectly in the form of public institutions that in one way or another shape the opportunities and possibilities that are open to them?

In *A Theory of Justice*, Rawls thought he could justify political coercion of this sort by showing that it fairly secured the conditions under which all citizens had the opportunity to realize a certain neo-Kantian ideal of autonomy. But he later came to the view that what he called “reasonable pluralism”—the existence of a plurality of conflicting but reasonable comprehensive doctrines—was deeper than he had at first assumed and as such could no longer support the Kantian ideal of citizenship. By the time he delivered the lectures of *Political Liberalism* in 1980, Rawls had become convinced that there were many reasonable, cooperative citizens who rejected Kantian moral doctrines of autonomy and could not on that account be convicted of irrationality, malice,

---

132 This is implied by Rawls’s definition of public reason, which is “the reason of equal citizens who, as a collective body, exercise final political and coercive power over one another in enacting laws and in amending their constitution” (PL, 214).

133 For example, both the distribution of educational resources and the content of the curriculum have a significant impact on the capacity of citizens to take advantage of opportunities to improve their social or economic situation.

134 The ideal of autonomy of *Theory* is viewed by Rawls as a development of the Kantian doctrine. For example, Rawls suggests that “we think of the original position as in important ways similar to the point of view from which noumenal selves see the world. The parties qua noumenal selves have complete freedom to choose whatever principles they wish; but they also have a desire to express their nature as rational and equal members of the intelligible realm with precisely this liberty to choose, that is, as beings who can look at the world in this way and express this perspective in their life as members of society” (TJ, 225). Where Rawls departs from Kant is in his view that “it is obviously impossible to develop a substantive theory of justice founded solely on truths of logic and definition. The analysis of moral concepts and the a priori, however traditionally understood, is too slender a basis. Moral theory must be free to use contingent assumptions and general facts as it pleases. There is no other way to give an account of our considered judgments in reflective equilibrium” (TJ, 44).
or unreasonableness. Under these circumstances, political institutions and laws could not be justified to all reasonable citizens on the explicitly Kantian grounds advanced in *Theory*. As Rawls puts it in the introduction to *Political Liberalism*, “[w]hile autonomy as a moral value has had an important [sic] place in the history of democratic thought, it fails to satisfy the criterion of reciprocity required of reasonable political principles and cannot be part of a political conception of justice. Many [reasonable] citizens of faith reject moral autonomy as part of their way of life.”

The Rawls of *Political Liberalism* not only rejected the moral viability or legitimacy of any deep, Kantian justification of liberal democratic political structures; he rejected the moral viability or legitimacy of *any* similarly contestable, deep philosophical justification, on the grounds that it could reasonably be objected to by citizens as the rationale for political authority and coercion. This realization led Rawls in search of a new justificatory strategy, which has become known generically as “political liberalism.”

Political liberalism invites us to make an important distinction between the values and propositions that should, or could, count as reasons for fundamental policy and law on the one hand—these are called “political” values or reasons; and the values and propositions that, however persuasive in some specialized domain, say religion or metaphysics, should not or could not count as justificatory reasons for fundamental law and policy—these are called “nonpublic,” “philosophical,” or “metaphysical” values or reasons.136 Public

---

135 Introduction to PL, xlv.

136 The domain within which public reason ought to be respected is a matter for debate. Rawls does not deny its pertinence to the wider lawmaking process, but he only insists upon its application to constitutional essentials and matters of basic justice. See PL, 215: “Some will ask: why not say that all questions in regard to which citizens exercise their final and coercive political values? To answer: my aim is to consider first the strongest case where the political questions concern the most fundamental matters. If we should not honor the limits of public reason here, it would seem we need not honor them anywhere. Should they hold
reason, which is essentially a fund of “political” values that can have justificatory force in arguments about matters of basic justice, consists of “presently accepted general beliefs and forms of reasoning found in common sense, and the methods and conclusions of science when these are not controversial,”\textsuperscript{137} or “the plain truths now widely accepted, or available, to citizens generally.”\textsuperscript{138}

The distinction between public and nonpublic reasons is at the very heart of Rawls’s ideal of public reason. It finds expression both in Rawls’s test for the legitimacy of a democratic political regime—what he calls the “liberal principle of legitimacy”—and in the so-called “duty of civility” which is thought to govern citizens’ political interactions. According to the liberal principle of legitimacy, “our exercise of political power is proper and hence justifiable only when it is exercised in accordance with a constitution the essentials of which all citizens may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational.”\textsuperscript{139} It is by no means self-evident what Rawls means by a constitutional essential that citizens “may reasonably be expected to endorse in the light of principles and ideals acceptable to them as reasonable and rational,” and how that differs from a constitutional essential that citizens may \textit{not} “reasonably be expected to endorse.” However, it is clear that Rawls does believe (or hope) there is some fund of principles and ideals that reasonable and

\textsuperscript{137} PL, 224.

\textsuperscript{138} PL, 225.

\textsuperscript{139} PL, 217.
rational persons can agree upon. I will explore the notion of reasonable endorsability later in light of Rawls’s “burdens of judgment” (§3.1.2 and 3.2.2).

Following close on the heels of Rawls’s liberal principle of legitimacy comes the duty of civility, which (partially) specifies what citizens need to do in order to conform their actions to the liberal principle of legitimacy, that is, in order to ensure that the constitution and the day-to-day operations of the law are such that all citizens “may reasonably be expected to endorse [them] in the light of principles and ideals acceptable to them as reasonable and rational.” The duty of civility specifies at a very general level the sorts of reasons citizens may (and may not) draw upon when arguing for laws and policies that touch upon matters of basic justice and constitutional essentials. It imposes upon citizens “a moral, not a legal, duty…to be able to explain to one another on [certain] fundamental questions how the principles and policies they advocate and vote for can be supported by the political values of public reason. This duty also involves a willingness to listen to others and a fairmindedness in deciding when accommodations to their views should reasonably be made.”

According to Rawls, liberal political values, or the values of public reason, have three basic features: first, they are somehow already implicit in our political culture; second, they have a limited scope, applying exclusively to the political domain or the “basic structure” of society; and third, they are “freestanding” from or do not presuppose the truth or validity (or falsehood or invalidity) of any particular comprehensive doctrine. Political values, according to Rawls, fall into two categories: first, the “values

140 PL, 217.

141 PL, 11-15. See my discussion of the concept of a “political” conception in §2.2.4 above.
of political justice,” which are supposed to be reflected in the basic structure or fundamental institutions of society, in particular the State and the economy; and second, the “values of public reason,” which provide moral standards for political inquiry and deliberation among citizens.\footnote{PL, 24.}

The values of political justice, at least on Rawls’s account, include equal political and civil liberty, equality of opportunity, and social equality. More specifically, they include freedom of conscience and expression, freedom of the press, the right to a fair trial, and the general freedom to pursue one’s favored conception of the good within the limits of justice. The values of public reason, on the other hand, include the virtue of reasonableness, which involves general epistemic responsibility and competence (for example, a willingness to consider relevant evidence and observance of basic logical canons), a moral responsiveness to the interests and rights of others, and observance of the so-called duty of civility, i.e. a willingness to offer “political” arguments for conclusions about matters of fundamental law or basic justice. Whether all of these values actually deserve to be included among “the political values of public reason” is a matter of great dispute, and it is the last value mentioned above—observance of the duty of civility—whose “political” credentials I hope to bring into question shortly.

\section*{§3.1.2 Grounds of the Duty: Reciprocity}

The reason nonpublic or nonpolitical reasons cannot have justificatory power in the political domain, at least when it comes to matters of fundamental justice, is not that they are \emph{false} (indeed, Rawls goes to some length to avoid making any such controversial
claim\textsuperscript{143} but that they run afoul of the value of reciprocity. By this, Rawls means that if I imaginatively reverse roles or positions with another citizen, assuming only that the other citizen is reasonable in the ordinary, straightforward sense of responsive to reasons, respectful of the canons of logic, and willing to adopt fair compromises for the sake of political cooperation and stability, I will perceive that the other citizen, given his reasonable set of epistemic and moral commitments, simply could not be expected to see the moral-political force or relevance of the reasons I offer.\textsuperscript{144} As Rawls puts it, the “criterion of reciprocity” is specified by the proposition that “our exercise of political power is proper only when we sincerely believe that the reasons we offer for our political action may reasonably be accepted by other citizens as a justification of those actions.”\textsuperscript{145}

Although Rawls does not here spell out that reciprocity only extends to “reasonable persons,” it is clear from other passages that the principle of reciprocity only binds us when we are addressing “reasonable” persons. For example, Rawls explicitly denies that “unreasonable doctrines” (and by implication, unreasonable persons) need be

\textsuperscript{143} “We try, so far as we can, neither to assert nor to deny any particular comprehensive religious, philosophical, or moral view, or its associated theory of truth and the status of values. Since we assume each citizen to affirm some such view, we hope to make it possible for all to accept the political conception as true or reasonable from the standpoint of their own comprehensive view, whatever it may be” (PL, 150).

\textsuperscript{144} Unfortunately, Rawls is never quite this explicit about what he means by reciprocity and the notion of “reasonable acceptability.” However, I believe this is a plausible interpretation of reciprocity, given that Rawls himself identifies as the “criterion of reciprocity,” the proposition that “our exercise of political power is proper only when we sincerely believe that the reasons we offer for our political action may reasonably be accepted by other citizens as a justification of those actions” (introduction to PL, xlvi).

Clearly, reciprocity involves a sense of fairness, but fair terms of cooperation are precisely those that every participant “may reasonably accept, provided that everyone else likewise accepts them” (PL, 16). It thus turns out that a basic orientation towards \textit{reasonable acceptance by all} is at the heart of the virtue of reciprocity.

\textsuperscript{145} PL, xlvi.
accommodated by political liberalism; on the contrary, such doctrines are to be “contained...so that they do not undermine the unity and justice of society.”

It is clear from the above that reciprocity requires us to advance justifications for coercive policies that could be accepted by any reasonable person, without abandoning or radically revising her current comprehensive doctrine. But we still have to investigate what Rawls means by a “reasonable” person, and this is no trifling matter. For the way we define the “reasonable” person directly shapes the constituency to whom we must make ourselves accountable, in speech, for our exercise of power. To put it negatively, those who are excluded from the category of “reasonable persons” become, to that extent, politically invisible or irrelevant, and need not be extended the sort of “reciprocal” political treatment that is extended to “reasonable persons.” Since much of my argument against Rawls (in particular, my interpretation of his principle of reciprocity) turns on how we are to understand his notion of the “reasonable person,” I will later examine this question at some length. But for ease of exposition, for now we can accept Rawls’s twofold characterization of the reasonable person as someone who (a) is willing to offer and abide by fair terms of cooperation; and (b) recognizes the morally innocent sources of disagreement among citizens (the burdens of judgment) and their consequences for the use of public reason in the exercise of power.

In order to get at what Rawls’s reciprocity requirement amounts to, we need to steer clear of two errors: on the one hand, to interpret it in such a demanding way that the reasons I offer must be viewed by any reasonable person as true and validly yielding my conclusion; or on the other hand, to interpret reciprocity in such an undemanding way

---

146 PL, xx.
147 PL, 54.
that the reasons I offer need only be acknowledged as validly yielding my conclusions, irrespective of their actual (perceived) epistemic credentials or plausibility. The relation between argument and audience required by reciprocity falls somewhere between full agreement on the one hand and acknowledgement of an argument’s validity at a purely formal level on the other. Roughly, in order for an argument to pass the test of reciprocity, any reasonable person must be able to see that the argument is intelligible, minimally coherent, drawing on reasons that can generally be grasped and assented to by reasonable citizens without some special leap of faith or personal conversion, and are, at least prima facie, consistent with the freedom and equality of all. I need not believe the reasons offered me for such-and-such a law are in fact true or conclusive, in order to accept that the reasons are intelligible, politically relevant, and consistent with the freedom and equality of all. To take a current example, if someone argues that we should introduce the death penalty on the grounds that it would act as a serious deterrent to violent crime, that studies show it is in fact administered impartially, and so on, I may vehemently disagree with the substance of his claims yet acknowledge that the argument is politically relevant and does not place any significant strain on my own reasonable view of the world: it is an argument I can engage in a rational way because it is framed in terms that, even if contestable, are accessible to our shared human reason.\footnote{Although the term “accessibility” can be misleading, it does capture an important dimension of the reciprocity requirement. Even a reason whose application or role in an argument I ultimately reject may, in an important sense, be accessible to me: after all, people disagree all the time about political outcomes while acknowledging that the reasons motivating their adversaries do have some weight in political matters, and are not inconsistent, as such, with the freedom and equality of all.}

Before considering the grounds of Rawls’s reciprocity requirement, I would like to clarify the role of reciprocity in the argument for the duty of civility. First of all, like
the concept of autonomy, I should emphasize that Rawls’s interpretation of the notion of reciprocity is one among many possible interpretations. So, whenever I refer to “the principle of reciprocity,” unless I say differently, I am referring to Rawls’s peculiar interpretation of the principle, not to the notion of reciprocity in general. Secondly, on Rawls’s view, reciprocity is a very general normative constraint upon the justification of exercises of political power that significantly impede the autonomy of others. It insists upon the accessibility or availability of our reasons for policies and laws to others differently situated, but it does not specify in greater detail what class of reasons is accessible. The duty of civility, on the other hand, is a more detailed specification of the constraints upon political discourse that follow from the very general constraint of reciprocity. For example, it specifies a class of reasons, “public” or “political” reasons, that are supposed to meet the reciprocity test, and a class of reasons, “nonpublic” reasons or reasons derived from “comprehensive doctrines,” that do not meet the test. According to the duty of civility, we must offer or be prepared to offer “political” or public reasons\textsuperscript{149} for our collective exercise of political power on matters of basic justice. The duty of civility, in short, is a concrete specification of the more general principle of reciprocity.

The reciprocity requirement and the associated duty of civility are motivated by two basic components of political liberalism: first, a moral view that emphasizes the importance of rational consent to laws that coerce citizens, whether directly or by structuring the opportunities and resources available to them; and second, an epistemological view offering a distinctive account of the nature of moral disagreement and its sources. I will begin by summing up the moral view as briefly as possible, and

\textsuperscript{149} See §2.2.5 above on the distinction between public and nonpublic reasons.
then outline Rawls’s account of moral disagreement. Rawls’s moral theory and (partial) epistemology jointly underpin the reciprocity requirement which is spelt out, if you will, by the duty of civility.

Let us start with the moral view: Rawls believes that coercive laws must be justified in terms that are rationally accessible to citizens (in a sense that will become clearer in light of his moral epistemology). What would be wrong with coercing a reasonable person based on reasons that he cannot endorse? In other words, why should a citizen confine his arguments for law or public policy to reasons that any reasonable person could endorse? The answer lies in the equal moral status possessed by persons, not as noumenal selves or children of God, but as citizens. Citizens are thought to possess an equal moral status, which can be cashed out in terms of an equal right to the opportunity to develop and exercise their two moral powers. The two moral powers are, first, the capacity to hold and revise a conception of the good; and second, the capacity to have and act upon a sense of justice. Citizens’ freedom consists precisely in their relatively unimpeded capacity to develop and exercise the two moral powers. Their equality consists of their right to an equal opportunity to develop and exercise their moral powers.\(^{150}\)

The basic idea is this: since each of us is free and equal, no one of us has any right to dispose of another or wield control over another’s life (special circumstances aside\(^{151}\)) without providing grounds for that intervention that are accessible to the other, i.e.

\(^{150}\) It is important to note that Rawls views this account of liberty and equality as specifying the conditions of political autonomy, not moral autonomy as understood by any particular comprehensive doctrine, say Kantian morality or natural law theory (PL, xliiv-xlv).

\(^{151}\) For example, the relation between adults and children, as well as between adults and the mentally incompetent, are not relations between equally competent and responsible agents, and as such, respect here takes on a different hue.
grounds that are not only intelligible, but cogent and capable in principle of being accepted by the other without stretching his current belief system to a breaking point. In other words, I owe you a justification for impeding your freedom to pursue your life goals, and not just any justification, but one you could reasonably and voluntarily accept. Otherwise, I am showing scant regard for your moral status as free and equal to me, and it will appear, from your standpoint, that I am just acting on reasons that I happen to believe (but have no weight for you), rather than appealing to our shared reason, or common sense. I might as well say to you, “You really ought to obey this law, because I believe it’s good for you to do so.”\textsuperscript{152} In short, the duty of civility reflects the requirement to respect the political autonomy, i.e. the freedom and equality, of other citizens, by justifying laws on grounds they can at least view as legitimate, understandable, and reasonable, even if mistaken in the case at hand.

Not surprisingly, considering that moral-political justification is at the heart of political liberalism, there is an epistemology, however rudimentary, at work in Rawls’s application of the test of reciprocity to political argumentation. Understanding that epistemology is essential in order to understand what the duty of civility and the liberal principle of legitimacy in fact require.\textsuperscript{153} The impetus for Rawls’s epistemology is of course not primarily theoretical but practical and political: Rawls is deeply impressed and

\begin{footnotesize}
\begin{enumerate}
\item[152] See PL, 247: “In recognizing others’ comprehensive views as reasonable, citizens also recognize that, in the absence of a public basis of establishing the truth of their beliefs, to insist on their comprehensive view must be seen by others as their insisting on their own beliefs. If we do so insist, others in self-defense can oppose us as using upon them unreasonable force.” For a neat summary of the broadly Rawlsian understanding of the relation between respect and public reason, see Eberle, Religious Convictions in Liberal Politics, 52-54.
\item[153] One interesting attempt to outline a full-blown epistemology of public reason, that initially takes its bearings from Rawls but rejects Rawls’s philosophical minimalism, is Gerald F. Gaus, Justificatory Liberalism: An Essay on Epistemology and Political Theory (Oxford and New York: Oxford University Press, 1996).
\end{enumerate}
\end{footnotesize}
concerned by the marked differences among apparently benign and conscientious inquirers in political and moral matters, since it seems to bring into question the liberal hope of providing a justification of the political order that is more or less satisfactory to all reasonable citizens.\(^{154}\) In other words, the wide divergence among apparently reasonable and rational persons on moral matters seems on its face to jeopardize the liberal hope of grounding the political order in citizens’ shared reason, or, if you will, their shared “considered moral judgments.”\(^{155}\) How can such a hope be revived in a pluralistic society? The answer, at least for Rawls, lies in the special claims of his epistemology which form the basis of an important moral argument. First, he offers an explanation of moral divergence, in terms of what he calls the “burdens of judgment”—namely, those factors, such as differing life experience and the difficulty of applying general principles to particular cases, that move people to draw different moral conclusions even though they are all basically conscientious, reasonable, and rational.

\(^{154}\) The differences that consistently arise in political and moral argumentation are also noted by Hobbes, and he too attempts to ground the political order in ways that will be persuasive to all, notwithstanding these differences. But whereas Rawls appeals to “shared truths” or “intuitions” of our political culture, Hobbes appeals primarily to the passions, and most fundamentally, the passion for self-preservation. See Hobbes’s *Epistle Dedicatory* to *De Cive*, xv: “From the two principal parts of our nature, Reason and Passion, have proceeded two kinds of learning, mathematical and dogmatical. The former is free from controversies and dispute, because it consisteth in comparing figures and motion only; in which things truth and the interest of men oppose not each other. But in the later there is nothing not disputable, because it compareth men, and meddleth with their right and profit; in which, as oft as reason is against a man, so oft will a man be against reason. And from hence it cometh, that they that have written of justice and policy in general, do all invade each other, and themselves, with contradiction. To reduce this doctrine to the rules and infallibility of reason, there is no way, but first to put such principles down for a foundation, as passions not mistrusting, may not seek to displace; and afterward to build thereon the truth of cases in the law of nature (which hitherto have been built in the air) by degrees, till the whole be inexpugnable.”

\(^{155}\) Rawls embraces a moderately optimistic anthropology, which insists that “[w]e must start with the assumption that a reasonably just political society is possible, and for it to be possible, human beings must have a moral nature, not of course a perfect such nature, yet one that can understand, act on, and be sufficiently moved by a reasonable political conception of right and justice to support a society guided by its ideals and principles” (lxii).
inquirers. Reasons infected by the burdens of judgment pertain to “non-public” reason; to be more precise, they are non-public with respect to the society of this or that liberal democracy: examples include religious claims about the revealed nature of Scripture; claims regarding the comparative merits of intellectual and political pursuits; and the claim that autonomy is the most important feature of a morally admirable human life. Second, he claims (or at least implicitly claims) that not all significant moral judgments are infected by the burdens of judgment: at least some reasons will be plausible to all in light of our shared human reason, the burdens of judgment notwithstanding. These reasons pertain to what Rawls calls the “political values” of “public reason,” and include values such as freedom of religion, freedom of speech, rule of law, and national security. Third, Rawls claims that a person who falls on the wrong side of an argument, where the matter is subject to reasonable disagreement because of the burdens of judgment, cannot be considered morally blameworthy for his mistaken beliefs; and that it would furthermore be unreasonable, or out-and-out unjust, to coerce that person for reasons that are vulnerable to the burdens of judgment. Fourth, Rawls argues that fundamental law must be grounded in reasons that escape the burdens of judgment, those “political values” that are “public” or available to any reasonable person. And this fourth claim presupposes, as Rawls recognizes, that “political values” are at least sufficiently

---

156 For a more detailed explanation of the burdens of judgment, see PL, 56-57. The burdens of judgment include conflicting empirical evidence, vagueness and indeterminacy of concepts with respect to hard cases, diversity of life experience, and conflicting normative considerations.

157 Rawls is not quite this explicit in his argument, but it seems to me to be the only plausible reading of his claim that “reasonable persons see that the burdens of judgment set limits on what can be reasonably justified to others...It is unreasonable for us to use political power, should we possess it, or share it, with others, to repress comprehensive views that are not unreasonable” (PL, 61).
determinate and extensive that they can yield reasonably well-grounded conclusions on all or most important matters of fundamental law or basic justice.\textsuperscript{158}

§3.2 A Rawlsian Argument against the Duty of Civility

Having set out in its essentials the content and rationale of the duty of civility, I now want to show that there is a powerful Rawlsian argument against the duty. The argument has three basic steps: (§3.2.1) first, I argue that the duty of civility, in light of (a) its pivotal role in the lawmaking process, and (b) the social sanctions associated with non-compliance, ought itself be justifiable to citizens in the same way as a fundamental law. (§3.2.2) Second, I argue that Rawls must justify the duty of civility to citizens who are reasonable in a relatively “thin” or philosophically non-partisan sense of the term, and that any attempt by Rawls to build acceptance of the duty into the very definition of reasonableness is inconsistent with other central elements of the reasonable (§3.2.3). Third, I show that the duty of civility cannot be justified to “thinly” reasonable citizens, given its epistemological and moral presuppositions, and therefore is not morally credible as a public guide for political deliberation.

§3.2.1 Should We Subject the Duty of Civility to the Test of Reciprocity?

The most fundamental grounding for the duty of civility is the protection of the two moral powers (jointly referred to as “political autonomy”) against political-moral violations. The duty of civility rests on the reciprocity requirement, which aims to protect

\textsuperscript{158} As Rawls puts it, “the substantive content and the guidelines of inquiry of a political conception” should be “complete,” meaning that “the values specified by that conception can be suitably balanced or combined, or otherwise united, as the case may be, so that those values alone give a reasonable public answer to all, or to nearly all, questions involving the constitutional essentials and basic questions of justice” (PL, 225).
citizens’ autonomy by insisting that political power should have a justification that appeals to our shared reason and could be reasonably accepted by all. The most obvious form of political power is direct enactments of law and policy, so these obviously fall under the reciprocity requirement: they need to pass some test of political justification. But what about other more indirect or subtle forms of political power? For example, what about the moral, not legal, rules that structure and constrain the *justification* of political power? As I will suggest shortly, there does not appear to be any reason to exempt these rules of justification from the reciprocity requirement, or consider them any less an exercise of political power than legal statutes or enactments themselves.

The duty of civility is just this sort of rule: it is a rule that, insofar as it is widely accepted (or at least accepted among the most influential sections of society, such as the mass media), structures justificatory political discourse by constraining or filtering the justification of political power. Of course, the cost of disregarding or violating the rule is not imprisonment or loss of property, but there are at least two reasons for believing that such a rule can have a serious impact on individuals’ political autonomy: first, if the duty of civility is widely recognized, violators will feel the force of social disapproval, and this sanction will weigh on them and put pressure on them to conform in ways comparable to the operation of legal sanctions, even when their comprehensive doctrines or conceptions of the good mandate that the rule be violated. Second, insofar as the duty of civility is 

---

159 One of the most famous descriptions of the subtle and powerful effects of non-legal, social rules on people’s lives is to be found in John Stuart Mill, *On Liberty* [1859], ch. 1, pp. 8-9, in *On Liberty and Other Essays* (Oxford: Oxford University Press, 1991), ed. John Gray: “Like other tyrannies, the tyranny of the majority was at first, and is still vulgarly, held in dread, chiefly as operating through the acts of the public authorities. But reflecting persons perceived that when society is itself the tyrant—society collectively, over the separate individuals who compose it—its means of tyrannizing are not restricted to the acts which it may do by the hands of its political functionaries. Society can and does execute its own mandates: and if it issues wrong mandates instead of right, or any mandates at all in things with which it ought not to meddle, it practises a social tyranny more formidable than many kinds of political oppression, since, though not
successful in structuring political discourse, those who reject it or whose conception of the good presupposes its rejection, will in general find that the prevailing pattern of justification for law makes it very difficult for them to contribute to that justification in effective and credible ways without disregarding values or beliefs they view as essential to a sound political argument. Thus, the laws they live under will be justified in ways they find they themselves cannot wholeheartedly participate in or endorse.\footnote{For examples, see my discussion of the implications of Rawlsian public reason for some Kantians, natural lawyers, Millians, and Catholics in §3.3.3 below.}

Because of the central, indeed water-shed, role played by the duty of civility in the process of political justification, and because it is not just a superogatory precept but a moral requirement of citizenship, it is tantamount to an exercise of political power and thus must be justified to citizens according to the same standards by which laws themselves are justified: namely, in terms that all citizens could reasonably be expected to endorse. We can unpack this by saying that in justifying the duty of civility, we must be able to imagine the justification having force for a citizen differently situated (the reciprocity test) in spite of, or over and above, the burdens of judgment. In short, the duty of civility not only rests on the reciprocity requirement: it must conform to that requirement itself on pain of forfeiting its political legitimacy.

§3.2.2 Who Are the “Reasonable” Citizens to Whom the Duty of Civility Must be Justified?

One matter that needs to be cleared up before my internal critique can proceed further is what Rawls means by “reasonable” person. Spelling out what is meant by

\footnote{usually upheld by such extreme penalties, it leaves fewer means of escape, penetrating more deeply into the details of life, and enslaving the soul itself.…}
reasonableness is crucial for my internal critique of Rawls, since it would appear that Rawls defines the reasonable person so narrowly that a reasonable person complies with the duty of civility simply by virtue of being reasonable. If that is the case, then Rawls is not required by his own theory to justify the duty of civility to anyone: it is simply taken for granted within his theory that any right-thinking, fair-minded or “reasonable” person in a liberal democratic culture will accept it. So it turns out that my internal critique of public reason stands or falls depending on how ecumenically we read Rawls’s notion of reasonableness.

Who are the “reasonable” persons to whom wielders of political power owe a justification, on Rawls’s account? Do they include people who might be skeptical of the duty of civility? Or are reasonable persons already committed, by definition, to accepting the duty? I will consider this question strictly from a Rawlsian perspective, with my internal argument in mind.\textsuperscript{161} If reasonable persons include dissenters from the duty of civility, then my internal argument against the duty might have some chance of getting off the ground, but if reasonable persons by definition accept the duty, then it would be hard to argue that the duty of civility is inaccessible to some reasonable persons, without repudiating Rawls’s own terminology and thus slipping into an external critique. It is therefore important for me to establish in advance of the main argument that the sense of “reasonable” I favor is firmly grounded within Rawls’s own theory rather than imposed from without. This is the task I set myself here.

The term “reasonable” appears in Rawls’s work in more forms and with more frequency than in most contemporary political writings. Unfortunately, the term is

\textsuperscript{161} In §3.3.2, I will consider which sense of the reasonable person we ought to assume for the purposes of a critique of public reason from integrity.
notoriously ambiguous and vulnerable to manipulation, and it is not always easy to
distinguish between inevitable semantic variation (depending, for example, on the object
being described as “reasonable”) on the one hand, and avoidable vagueness and
equivocation on the other. The two usages of “reasonable” that will concern us here are:
first, concerning comprehensive doctrines; and second, concerning persons. We will find
that these two usages are intimately linked, and that semantic variation in one application
of the term (the reasonableness of comprehensive doctrines) can yield variation in the
other application (the reasonableness of persons) and vice versa.

Unfortunately, Rawls says very little about what counts as a “reasonable” or
“unreasonable” comprehensive doctrine, and what he does say is internally inconsistent
and (at least partially) in tension with other parts of his theory. On the one hand, Rawls
asserts in the introduction to *Political Liberalism* that “political liberalism…supposes that
a reasonable comprehensive doctrine does not reject the essentials of a democratic
regime,”

thus giving some (albeit somewhat vague) normative bite to the concept. But
later in the book, he defines reasonableness with respect to comprehensive doctrines so
loosely that it could apply to comprehensive doctrines that on most interpretations,
including their own, “reject the essentials of a democratic regime.” On this second view,
a reasonable doctrine has three features: (a) it is “an exercise of theoretical reason,”
covering “the major religious, philosophical, and moral aspects of human life in a more
or less consistent and coherent manner”; (b) it is “an exercise of practical reason”; and (c)
it normally “belongs to, or draws upon, a tradition of thought and doctrine,” and “tends to
evolve slowly in the light of what, from its point of view, it sees as good and sufficient

---

162 PL, xviii.
reasons.” But this description of reasonable comprehensive doctrines, as Lief Wenar points out, would allow for anti-democratic doctrines including white supremacism and Islamic fundamentalism.

The first definition of reasonable doctrines, requiring that they accept the “essentials of a democratic regime,” is pretty vague, but at least it can cohere with Rawls’s account of the “reasonable person,” and helps us make sense of the distinction between simple pluralism (plurality of comprehensive doctrines) and reasonable pluralism (plurality of “reasonable” comprehensive doctrines) which lies at the very basis of Rawls’s shift to political liberalism. Since Rawls says a lot more about the “reasonable person” than about “reasonable comprehensive doctrines,” I favor Wenar’s strategy of reading the latter as a “comprehensive doctrine that can be affirmed by a reasonable person,” which makes “reasonable person” the controlling term in the definition.

We are thus driven back to the meaning of a “reasonable person.” Fortunately, Rawls does spell out in some detail what he has in mind. Wenar offers a more extensive interpretation, collecting elements of the “reasonable” scattered throughout PL. But for our purposes, it is sufficient to highlight the most fundamental aspects of reasonableness. Rawls offers two distinct interpretations of the reasonable person, one as part of his explanation of the burdens of judgment (i.e. those factors that differentially affect our moral judgments without thereby calling into question our bona fides or moral

---

163 PL, 59.
165 See Ibid., 36, fn. 6.
166 Ibid., 37.
credentials), and the other in his main account of the “reasonable person.” The two interpretations differ in that the latter incorporates more elements than the former. For convenience, I will call these Rawls’s thin and thick accounts respectively. Let’s begin with the thinner account of the reasonable—the account associated with the burdens of judgment.

One of the two “basic aspects” of reasonable citizenship (what I am calling the thick account of reasonableness), as we shall see shortly, is “the willingness to recognize the burdens of judgment and…their consequences for…public reason.” But the burdens of judgment are precisely the “sources, or causes, of disagreement between reasonable persons.” In order to avoid circularity, we should not assume here that Rawls has in mind the thick account of reasonableness which recognition of the burdens of judgment is just one part of. For otherwise, the burdens of judgment would be the sources of disagreement between persons who recognize the burdens of judgment. Besides, the text seems to point us to a notion of reasonableness that is distinct, though closely related, to that of the main account it is supporting:

Let’s say that reasonable disagreement is disagreement between reasonable persons: that is, between persons who have realized their two moral powers to a degree sufficient to be free and equal citizens in a constitutional regime, and who have an enduring desire to honor fair terms of cooperation and to be fully cooperating members of society. Given their moral powers, they share a common human reason, similar powers of thought and judgment: they can draw inferences, weigh evidence, and balance competing considerations…The idea of reasonable disagreement involves an account of the sources, or causes, of disagreement between reasonable persons so defined.169

167 PL, 54.
168 PL, 55 (emphasis added).
169 Ibid.
To sum up the (thin) account of reasonableness associated with the burdens of judgment, we could say that it requires persons to be sufficiently competent reasoners (the epistemic or rational component) and sufficiently responsive to the demands of justice or the interests of others (the moral component) to be free and equal citizens. This does not appear to be a hugely demanding benchmark of reasonableness. Indeed, I believe it would be hard for anyone, at least in a Western democracy, to credibly bring it into question.

The “thick” account of the reasonable is offered in section 1 of Lecture 2. This section is entitled “The Reasonable and the Rational” and it is the most extensive and, judging by the section title, the most self-conscious account of a reasonable person that Rawls has to offer. I am assuming that this captures more adequately Rawls’s considered views on reasonable citizenship than the sparser, nested account of reasonableness discussed above. In the thick account, Rawls points to two “basic aspects” of the reasonable, both of which are “virtues of persons.” ¹⁷⁰ The reasonable person evinces a willingness to “propose fair terms of cooperation and to abide by them provided others do” as well as to “recognize the burdens of judgment and to accept their consequences for the use of public reason in directing the legitimate exercise of political power.” ¹⁷¹ On Rawls’s interpretation of fairness, “[f]or these terms to be fair terms, citizens offering them must reasonably think that those citizens to whom such terms are offered might also reasonably accept them” ¹⁷²; while on Rawls’s interpretation of the “consequences of the burdens of judgment for the use of public reason,” the permanence and inevitability of

¹⁷⁰ PL, 48.
¹⁷¹ PL, 54.
¹⁷² PL, xlv.
morally innocent philosophical and religious disagreement precludes fair-minded citizens from enacting fundamental laws and policies based on their comprehensive doctrines. Thus, Rawls’s thick concept of the reasonable effectively includes within it, via Rawls’s interpretation of “fair terms of cooperation,” the duty of civility, or the duty to offer others reasons for fundamental policies and laws that are accessible to their shared reason—reasons they could accept in spite of their differing comprehensive doctrines. On this interpretation (what I am calling the “thick” concept of reasonableness), it is just part of the definition of a reasonable person that she accepts and is prepared to abide by the duty of civility.

It might seem, then, that however questionable the imposition of the duty of civility may be, the duty can be deduced from Rawls’s account of the reasonable person and as such is not vulnerable to an internal critique. This definitional victory, however, proves upon closer inspection to be an illegitimate one. Even if we accept Rawls’s general description of reasonableness, viz. the willingness to offer fair terms of cooperation and recognize the burdens of judgment and their consequences for public reason, this does not grant Rawls a carte blanche to interpret or apply these general requirements in any way he wishes. From a strictly internal Rawlsian perspective, any interpretation of reasonableness must meet a test of theoretical coherence, or coherence with other important commitments of political liberalism. Imagine, for example, if Rawls had defined a reasonable person as somebody who, among other things, accepts that the highest life for man is the life of contemplation. This definition of the reasonable would clearly conflict with other central aspects of the reasonable, in particular the need to recognize the burdens of judgment and their consequences for the use of public reason.
Thus, Rawls is not entitled to take for granted any specification of reasonableness he pleases, including acceptance of the duty of civility: his specification of the reasonable must conform to other central requirements of reasonableness on pain of internal contradiction. One of these is sensitivity to the burdens of judgment, which requires that the grounds for coercion be accessible to thinly reasonable persons. Indeed, I will argue in the next section that imposing the duty of civility fails to respond to the burdens-of-judgment aspect of reasonableness.

What I have shown so far is not that Rawls’s thick conception of reasonableness is flawed or wrong-headed, but that (a) it imposes the duty of civility; and (b) insofar as this duty involves an exercise of power, as I have already argued (§3.2.1), its acceptance cannot be taken for granted by definitional fiat—on the contrary, arguments for its imposition must be sensitive to the burdens of judgment, or accessible to “thinly” reasonable persons in spite of the burdens of judgment. The next stage of the argument will be to show that the duty of civility cannot in fact be justified to (thinly) reasonable persons without disregarding the burdens of judgment. To that extent, neither the duty, nor the thick conception of the reasonable that embodies it, is morally legitimate or reasonable from a Rawlsian perspective.

§3.2.3 Does the Duty of Civility Pass the Test of Reciprocity?

If I am right, then the duty of civility is so central to the lawmaking process that it must itself pass Rawls’s test of reciprocity, i.e. be justifiable in terms (thinly) reasonable persons could accept without abandoning their comprehensive doctrines. But does it? Perhaps if it were framed vaguely in terms of publicity or accessibility of reasons, it
could. But what the duty of civility in fact presupposes is a distinctive interpretation of
notions such as reciprocity, publicity and accessibility, that is driven by (a) a neo-Kantian
or egalitarian moral epistemology associated with Rawls’s account of the burdens of
judgment; and (b) a contractarian political morality that emphasizes universal consent to
or universal justifiability of coercive rules as a virtually overriding moral requirement.
The burden of this stage of the argument is to show that these presuppositions of the duty
of civility, including Rawls’s principle of reciprocity, cannot satisfy the reciprocity
requirement. My argument issues in one simple claim: that *Rawls’s duty of civility is
fundamentally incoherent, because the reciprocity principle underlying it fails to conform
to itself, i.e. cannot be justified to epistemically and morally responsible (‘thinly
reasonable’) citizens, and is thus an illegitimate exercise of public reason.*

What lends a certain appeal to the duty of civility, at least initially, is (a) the fact
that people do disagree on political and moral matters without any evident malice,
irrationality or bad faith and (b) the almost universally accepted norm (at least in Western
cultures) that political authority ought to be exercised in ways that can be rationally
justified to all concerned. When there is deep disagreement and no way to resolve it
rationally, we are naturally reluctant to consider one party to the disagreement justified in
imposing its opinion on the other. This is most obvious in the case of religious doctrines:
what sort of argument might we use, for example, to show everyone that the Catholic
Church is the one true church and the surest path to salvation? Is this a proposition that
can be settled in a rational, public manner, based on reasons that are somehow accessible
to all reasonable and rational persons? Most of us think not. And in the absence of any
public rational settlement of the issue, any attempt to *enforce* one side of the dispute
politically seems arbitrary and violative of the right of citizens to live their lives according to their own lights or at least to (have a reasonable opportunity to) be reconciled through rational principles with the rules under which they live.\textsuperscript{173}

The problem is that what seems to be a universally recognized fact stands in need of interpretation. There is no uniquely authoritative or “common-sensical” interpretation of the fact of divergence in political and moral judgment, and certainly no uniquely authoritative or “common-sensical” interpretation of the political implications of that fact. For in order to interpret conflicts in moral judgment and draw out their political implications, we require (a) a moral epistemology—however rudimentary; (b) a political morality or theory of political justification; and (c) a sociology or interpretation of our social practices, however basic. And the Rawlsian hope that all (or indeed any) of these can somehow overcome or escape the burdens of judgment, and thus meet the reciprocity requirement, is an empty one, as we shall see.

What, then, are the specific epistemological, sociological, and moral assumptions underlying the duty of civility? Let us start with Rawls’s moral epistemology: notice that he goes to great lengths to preserve the bona fides and good intentions of the various parties to political disputes. When a deep moral dispute arises among ostensibly cooperative citizens, Rawls favors an epistemological explanation for the divergence that does not in any way undermine the moral credentials or competency of the various parties

\textsuperscript{173} This political ideal is reflected in the third role Rawls identifies for political philosophy in John Rawls, \textit{Justice as Fairness: A Restatement}, ed. Erin Kelly (Cambridge, Massachusetts and London, England: The Belknap Press of Harvard University Press, 2001): “political philosophy may try to calm out frustration and rage against our society and its history by showing us the way in which its institutions, when properly understood from a philosophical point of view, are rational, and developed over time as they did to attain their present, rational form” (3). Cf. Jeremy Waldron, who traces the social contract back to a view he believes “most liberals…share: that the social order must be one that can be justified to the people who have to live under it” (“Theoretical Foundations of Liberalism,” in Jeremy Waldron, \textit{Liberal Rights: Collected Papers 1981-1991} (Cambridge: Cambridge University Press, 1993), 57-58.
to the dispute. Where an Aristotelian or Thomist might look to passion, vice, and rationalization (and not just the burdens of judgment) as an explanation for moral disagreement, Rawls privileges the “burdens of judgment” and seems decidedly reluctant to give a major explanatory role to moral defectiveness, bad faith, self-interest, or rationalization.\textsuperscript{174} Where an Aristotelian assumes that some moral agents are wiser and more competent moral judges than others on account of their superior virtue, Rawls assumes with most neo-Kantians that most of us are equally competent moral judges whose moral qualifications cannot credibly be brought into question, unless perhaps we deliberately flout universally recognized moral rules. In other words, one of the premises driving Rawls’s account of moral divergence is a kind of moral egalitarianism that is distinctive of neo-Kantian morality.\textsuperscript{175}

In order to get from a presumption of good faith and equal competency in moral disagreement to the primacy of political over comprehensive justifications, Rawls needs to rely, at least implicitly, upon a reasonably contestable contractarian political morality in addition to some controversial claims of fact about our society. Start with the controversial sociological claims: Rawls claims that there is, \textit{in fact}, a commonly-

\textsuperscript{174} Thus, Rawls insists that “[t]he account of [the sources, or causes, of disagreement among reasonable persons] must be such that it is fully compatible with, and so does not impugn, the reasonableness of those who disagree” (PL, 55). But even if we grant (as I readily do) that many moral disagreements are of this kind, it does not follow that all are, or that the most salient issues in political morality are of this kind. This would need to be argued, yet Rawls does not offer us reasons for this position. Instead, he simply asserts that this is the “right kind” of explanation for moral disagreement in a democratic culture (ibid.).

\textsuperscript{175} The alternative view of ethics, which distinguishes more from less mature moral agents, is cogently put by Hursthouse, who points out that virtue ethics (unlike Kantian ethics) “takes seriously Aristotle’s point that moral knowledge, unlike mathematical knowledge, cannot be acquired merely by attending lectures and is not characteristically to be found in people too young to have much experience of life. A normative ethics should not aim to provide a decision procedure which any reasonably clever adolescent could apply” (Rosalind Hursthouse, \textit{On Virtue Ethics} (Oxford and New York: Oxford University Press, 1999), 18). For an Aristotelian/Thomistic treatment of moral divergence, see Robert P. George, \textit{Making Men Moral} (Oxford: Clarendon Press, 1993).
accepted political morality, or at least the elements of a political morality, already implicit in our society’s political traditions and practices, significant or ‘thick’ enough to do the work of public reason, i.e. to meaningfully adjudicate disputes that arise over basic justice and constitutional essentials; and that what he calls the “political values of public reason,” as specified in political liberalism, constitute a reasonable interpretation of this political morality that any reasonable person could endorse. But the practices of our democracy consistently belie this claim. Time and time again, constitutional disputes and disputes over basic justice are cast in terms of different fundamental moral considerations by different parties to the dispute, all claiming to espouse the same democratic heritage, none claiming to reject the traditional values of constitutional democracy such as liberty, equality, due process, and rule of law. For example, while ostensibly disputants all subscribe to the value of liberty, their interpretations of liberty diverge so dramatically that people seem to be talking past each other rather than engaging in a meaningful exchange of arguments.176 Examples of this abound, but we need only mention a few here: whether polygamy is sanctioned by freedom of religion; whether euthanasia is an act of compassion or of irresponsible killing; whether there is a constitutional right to, permission of, or prohibition of, same-sex marriage; whether school prayer is a violation of the first amendment; whether there is a constitutional or moral entitlement to, permission of, or prohibition of, abortion.177 Most of us affirm the same basic concepts

176 See Alasdair MacIntyre, *After Virtue: A Study in Moral Theory* (London: Duckworth, 1981), for a strong argument for the thesis that moral argumentation in contemporary Western societies such as the United States is frequently futile and “interminable.”

177 George Klosko undertakes an empirical study of American public opinion to determine whether, or to what extent, there is evidence of an “overlapping consensus” on political values, and not surprisingly, he finds that though there is a rough consensus (somewhere around 70%) on constitutional values stated in the abstract, and on certain procedural values such as due process and rule of law, the consensus quickly
(liberty, equality, etc.) at a certain level of abstraction, but espouse substantially different interpretations of those concepts and what they entail in practice. In short, the sociological preconditions for public reason, at least as Rawls envisages it, are at best highly questionable—though I would be inclined to say they are simply unsustainable. For the purposes of my argument, all I need to show is that they are highly questionable.

Aside from some controversial claims of fact (a basic sociology, if you will), the primacy of the political is grounded in at least one crucial normative assumption associated with a contractarian political morality. This assumption is not always fully spelt out by Rawls, but deserves more attention than it receives. The assumption is that any minimally competent, cooperative, epistemically responsible, moral agent ought to be capable in principle of consenting to the laws that govern his conduct, where “in principle” does not mean “as a human being,” but in his present moral condition—assuming only participation in a liberal democratic political culture and some minimal moral requisites such as cooperativeness. And this is not just one desideratum to be weighed against many others, but central and primary in the justification of law. Political liberalism, in other words, places a premium on consent to law that is in keeping with a vision of society as a contract among equals. Where consent comes into conflict with other values or goods such as the moral ecology of society, unborn life, or the stability of family life, consent tends to win out, since it is very difficult to secure, or even imagine, the unanimous agreement of minimally reasonable people on those other values.

dissipates as soon as those values are applied to individual cases (George Klosko, Democratic Procedures and Liberal Consensus (Oxford: Oxford University Press, 2000)).

It was Charles Taylor, Modern Social Imaginaries (Durham and London: Duke University Press, 2004) which forcefully brought home to me the central and distinctive role of contract in the modern self-understanding. Cf. my discussion of political liberalism as a species of contractualism in §5.3.1 below.

For the notion of a “moral ecology” to be protected by law, see George, Making Men Moral.
Once we assume the centrality of consent to political justification, it then makes sense to seek out a category of reasons that are or can be persuasive to any reasonable person, reasons that can function as a resource for justifying law to all reasonable citizens: what Rawls calls “political” reasons. Furthermore, granting the centrality of consent to justification, it makes sense to withhold coercive measures to the extent that they fail to secure a suitably “political” justification.

I want to argue that the moral primacy or virtually overriding weight attached to reasonable consent by political liberalism is itself something that is subject to reasonable disagreement, and is eminently susceptible to the burdens of judgment. In the event that I am right on this point, the reciprocity requirement grounding the duty of civility is itself reasonably contestable and therefore fails the reciprocity test. It is literally self-defeating.

In order to show that the moral primacy of consent presupposed by the duty of civility is subject to reasonable disagreement, I will show how this disagreement is subject to the burdens of judgment in a dispute among (thinly) reasonable persons. Once this is established, the duty of civility, insofar as it occupies a central and weighty place in the political justificatory process, can be seen to violate citizens’ political autonomy, and as such, is best dispensed with if we are to be faithful to Rawls’s own principles.

Reasonable—that is, thoughtful, epistemically and morally responsible—citizens may use their rational and moral powers to reach different conclusions, in good faith, about (i) how we are to understand consent (whether hypothetically, tacitly, actually, or otherwise) and (ii) the weight to be attached to consent vis-à-vis other moral desiderata in the justification of law and policy. There is nothing rationally self-evident or obvious about how to either (i) specify the desideratum of consent or (ii) weigh it against other
desiderata such as justice, human well-being, rational autonomy, etc. We might be inclined to agree with certain cases where social justice was furthered by acts that apparently overrode the consent of (some) reasonable persons, e.g. the American civil war and Brown v. Board of Education. However, it seems much less clear-cut whether or not we should override the consent of drug addicts (not guilty of other crimes such as theft), viewers of pornographic films, parents who educate their children in unconventional values, parents who fail to provide adequate parenting to their children, etc., for the sake of some other good, or for the sake of the well-being of the coerced. To the argument that these cases do not fall under constitutional essentials and basic justice and therefore are not affected by the requirement of reciprocity, I would reply that very great interests of persons are at stake in such cases, including personal autonomy and integration into society. To the extent that such cases (or other cases like them) implicate constitutional essentials and basic justice, there is no reason from Rawls’s perspective to exempt arguments about them from the duty of civility.

To demand that the principles settling these problems meet a reciprocity requirement is effectively to take sides in a dispute among reasonable persons about a complex issue that is subject to the burdens of judgment. Imposing a reciprocity requirement prior to political discourse preemptively settles the problem of how much weight to attach to consent vis-à-vis other moral desiderata. But this is a problem subject to the burdens of judgment enumerated by Rawls himself: (i) different considerations, even assuming their relevance is accepted by all, may be weighted differently by different persons. The autonomy of the person does not obviously trump his physical and

psychological health and well-being, nor does it obviously trump common goods such as public order and the stability of family life, except in the eyes of some; (ii) our moral and political judgments are influenced by our total life experience, which necessarily differs across persons. How we judge the problem of consent is obviously affected by the political and religious beliefs of our parents and friends, and the ways in which we have experienced our own freedom and the freedom of others in our lives, the goods or evils we have seen flowing from acts overriding consent, etc. (iii) Normative considerations may be advanced in favor of both sides of an issue, producing a conflict of values difficult to mediate: clearly, there is a normative argument for weighing consent very heavily in our political judgments deriving from the value of political autonomy; but there is also an argument for weighing the common good and the well-being of individuals more heavily than autonomy, derived from a more substantive or positive ideal of freedom. (iv) A hard choice between many possible goods must be made in social and political policy-making—there is a limited social space which cannot instantiate the full range of values. If we reflect upon what kind of social goods we wish to instantiate, we are forced to choose between some goods associated with individual freedom to experiment with one’s life free from coercion, and other goods associated with a social environment structured by shared expectations of non-destructive behavior (e.g. the expectation that members of the society ought not use addictive drugs).

I have argued that the duty of civility, given the significant psychological and social constraint it imposes on citizens’ participation in the political process and on the laws and policies emerging from that process, must itself be justified in a way that meets Rawls’s reciprocity test, on pain of violating citizens’ political autonomy. I have argued
that the reciprocity requirement or test ought to be interpreted, from a Rawlsian perspective, as imposing upon citizens the duty to justify legal limitations of others’ autonomy in terms any thinly reasonable person (i.e. anyone both epistemically responsible and competent, and prepared to offer and abide by fair terms of cooperation) could accept, notwithstanding differences in their reasonable comprehensive doctrines. But as I have shown, the reciprocity requirement itself, which forms a central part of the justification for the duty of civility, attaches a high and often overriding priority to consent. Given that the weight attached to consent as a moral-political desideratum is a matter over which (thinly) reasonable persons can reasonably disagree, the justification for the duty of civility, viz. the reciprocity requirement, is literally self-defeating, i.e. it fails its own test of reciprocity. In short, the duty of civility and its underlying reciprocity requirement are not justified in terms any reasonable person could accept, and thus violate citizens’ autonomy.

§3.3 The Critique from Integrity

Having investigated the internal workings of Rawls’s ideal of citizenship and found it to be ultimately self-defeating in its reciprocity requirement (as expressed in the duty of civility), I now turn to consider more carefully the implications of the duty of civility for the ethical integrity of citizens.\footnote{The internal argument certainly touches on integrity when it considers the conflict between citizens’ ethical views and public reason, but I now want to bring this issue to the fore.} One of the questions motivating my exploration of the Rawlsian ideal of citizenship is, whether it is possible for a liberal democratic ideal of citizenship to accommodate citizens’ efforts to live lives of integrity, or better, to what extent is such accommodation a realistic goal? Of course, even if
Rawls’s ideal fails to accommodate integrity, that does not imply that liberal democracy is incompatible with integrity. However, given the pervasive influence of political liberalism over contemporary liberal theory, and its paradigmatic status, I hope that understanding its implications for integrity will provide some sense of direction to any project (such as mine) that aims to forge an ideal of citizenship friendly to integrity yet tailored to a liberal democratic regime.

The basic claim I will defend here is that political liberalism, especially in its moral strictures on political speech, prevents reasonable citizens\textsuperscript{182} from acting upon their reasonably and conscientiously held beliefs in the public arena. Furthermore, in the absence of any compelling public purpose for doing so, this seems contrary to the liberal presumption in favor of freedom. I am building off and expanding upon Nicholas Wolterstorff’s critique of Rawlsian public reason, which points out that for many Christians, the notion of subordinating the will of God as revealed, say, in Scripture, to any principles in the public domain is anathema to their way of life and their most central moral convictions. For such “citizens of faith,” to use Rawls’s term, any effort to set aside “political values” as “freestanding” from their religious beliefs and as having normally overriding weight in the public domain,\textsuperscript{183} would be to lead a double life: it would be to require religious believers to consider religious values of overriding weight

\textsuperscript{182} As I will suggest in §3.3.2, there are compelling reasons, both within political liberalism and independently of it, for favoring the less demanding or “thinner” sense of reasonable rather than Rawls’s “thicker” sense.

\textsuperscript{183} “Given the existence of a reasonably well-ordered constitutional regime,” according to Rawls, “two points are central to political liberalism. First, questions about constitutional essentials and matters of basic justice are so far as possible to be settled by appeal to political values alone. Second, again with respect to those same fundamental questions, the political values expressed by its principles and ideals normally have sufficient weight to override all other values that may come in conflict with them” (PL, 137-138). Later, Rawls suggests that one of the conditions to be satisfied if we are to “honor public reason” is that “we give very great and normally overriding weight to the ideal it prescribes” (241).
in the privacy of the church or home, while considering them only of secondary importance in the public square.\textsuperscript{184}

§3.3.1 Integrationist and Non-Integrationist Citizens

What I propose to do here is expand Wolterstorff’s argument by considering the impact of public reason not only on religious citizens but on citizens who seek to pursue lives of integrity, whatever their attitude towards religion or Christianity.\textsuperscript{185} A good place to start is with a rough description of the citizen who seeks to live with integrity—let us call her the “integrationist” citizen for short. For now a fairly rudimentary interpretation should suffice (it will not be until chapter 4 that I develop a fuller account of integrity). First of all, think of an integrity requirement as the requirement that one’s diverse values, beliefs and actions hang together to form a coherent whole, such that no two values, beliefs, or actions, however diverse the relevant spheres of activity, contradict each other.\textsuperscript{186} For example, if you believe that dishonest or deceptive bargaining is wrong in

\textsuperscript{184} This is a gloss of the argument Wolterstorff makes in Audi and Wolterstorff, \textit{Religion in the Public Square}. This position does not entail that the claims of political authority are irrelevant or carry no weight: it only entails that the Christian will always follow God’s will according to his best judgment, both when God’s will sanctions obedience to political authority and when God’s will calls for resistance to or non-cooperation with authority. The same goes for social and cultural norms, including the norms of public reason: a Christian must set them aside wherever he deems that they conflict with God’s will.

\textsuperscript{185} Phil Quinn argues that the concern with integrity is neither characteristic of \textit{all} Christians, nor exclusive to them, in Quinn, “Can Good Christians Be Good Liberals? On the Ambivalent Answer of Nicholas Wolterstorff.”

\textsuperscript{186} Of course, in many cases what constitutes and does not constitute a “contradiction” is a highly contestable matter. For example, Rawls might claim that applying criteria of right and wrong in public affairs that differ from those applied within one’s church or family does not necessarily constitute a contradiction. Indeed, in certain cases this may be so. However, the values one holds and applies in non-public scenarios may have implications for public decisions and behavior, and denying or ignoring those implications would constitute a contradiction within a person’s moral life. For example, if I treat the poor and the ill with compassion and care in my private dealings, and then treat them with “strict justice” in a more political setting, I have to explain how these two attitudes, coming from the \textit{same person}, can be reconciled.
most of your human relations, you may find that the principle or disposition leading you to that belief also rules out manipulative bargaining in the workplace, or government, or marketplace. Or suppose you believe that racial abuse is an affront to human dignity: you may feel compelled by this belief at least not to turn a blind eye to instances of racial abuse, whether in your home, or in your workplace, even if your job prospects will be jeopardized or your most powerful and influential friends at work are instigating the abuse. Finally, to take a more controversial example, if you believe on ethical and/or sociological grounds that heterosexual marriage advances important social and moral goods including overall human flourishing, while same-sex marriage is likely to undermine them, you may feel compelled to oppose same-sex marriage in the political arena so that your considered judgments about marriage and family life are consequential not only for your familial and community dealings, but also for your reflections as a citizen on law and the political common good. Clearly, a fuller argument spelling out missing premises would be needed to demonstrate, or even make a strong case, that in each case the person involved is right about the implications of integrity. However, I am less concerned to demonstrate precisely what integrity requires than to illustrate roughly how integrity often amounts to a responsiveness of the agent to certain ethical or moral principles in all activities however diverse.187

187 Notice that a responsiveness of the agent to certain ethical or moral concerns does not imply that all of the relevant concerns can be fully met, or that full integrity is always a live option. My argument does not mean to exclude (nor does it mean to affirm) the possibility of tragic moral dilemmas, or inherently irresolvable moral conflicts that are not due to the fault of the agent. The locus classicus for the problem of moral dilemmas is Walter Sinnott-Armstrong, Moral Dilemmas (Ann Arbor, Michigan: University Microfilms International, 1982). Other approaches include Bernard Williams, "Consistency and Realism," Aristotelian Society Suppl 40, no. 1-22 (1966); Bernard Williams, “Ethical Consistency,” in Problems of the Self: Philosophical Papers 1976-1972, ed. Bernard Williams (Cambridge: Cambridge University Press, 1973); and Ruth Marcus, "Moral Dilemmas and Consistency," Journal of Philosophy 80, no. 77 (1980).
Now, when I speak of an integrationist conception of the good or comprehensive doctrine, I mean a moral view that aspires towards overall unity, not only within a particular value domain (say, health, or leisure, or politics, or intimate relationships), but across all the value domains of one’s life. This desired unity is always a unity in the making, a regulative ideal so to speak, so that it is more a matter of making adjustments to one’s life to overcome value conflicts as they arise and are noticed, than actually achieving full unity at a given moment. The point to be stressed here is that unity is desired across rather than merely within value domains. As such, a person with an integrationist conception of the good will not be satisfied by the thought that the values of one domain, say politics, are simply irrelevant or tangential to another, say, familial relations or the spiritual life. For the integrationist, the moral life is largely a matter of attaining greater coherence among the diverse spheres of life, such that the values reflected in each sphere are consistent not only internally but with each other.\footnote{I am not suggesting that the criteria of consistency and contradiction are straightforward or easily discerned in all cases. But if our reasoning in, say, a political forum, is different in character to our reasoning about domestic matters, it always remains the practical reasoning of a single agent who (we hope) aims to live well overall. Granted, doing good has different concrete implications depending on the object and purposes of our action; but this does not entail that we cease to reason as a person with one view of the good (however provisional or messy) upon entering into this or that social role.}

Non-integrationist conceptions of the good, on the other hand, while they may require some degree of unity within a life, are much less demanding in this respect than integrationist conceptions. Thus, while they may demand that the diverse values within a particular sphere of life (say, politics or friendship) form a coherent whole, they do not demand or expect such unity among values across the different value domains or activities of life. They are capable of making strong distinctions between, say, the
political domain of values and the comprehensive or more personal domain, and are willing to treat some value domains as more or less autonomous from others.

§3.3.2 Reasonable vs. Unreasonable Integrationist Citizens

In order to pre-empt one possible rejoinder to my critique from integrity, I must stipulate in advance that the integrationist citizens with whom I am concerned here are reasonable in the sense of rational, fair-minded, cooperative, and responsible. They embody all the basic dimensions of Rawls’s “thin” sense of reasonableness, that is, the sense that is operative in his account of the burdens of judgment: they are “persons who have realized their two moral powers to a degree sufficient to be free and equal citizens in a constitutional regime, and who have an enduring desire to honor fair terms of cooperation and to be fully cooperating members of society. Given their moral powers, they share a common human reason, similar powers of thought and judgment: they can draw inferences, weigh evidence, and balance competing considerations.”189 We might surmise that reasonable citizens, so understood, do not have totalitarian impulses and do not want to overthrow liberal institutions and practices (though they are likely to favor various reforms). One would expect a reasonable citizen of a liberal democratic regime to support a broad panoply of rights and values associated with liberal democratic regimes, such as rule of law, freedom of religion, freedom of conscience, freedom of association, freedom of the press, the right to a fair trial, equality before the law, democratic procedures, and so on, though they may differ from each other and from non-integrationist citizens in their interpretation and application of such values.

189 PL, 55.
So far, so good. If this account of reasonableness is correct, then the integrationist citizens I am concerned about, insofar as they are reasonable, are part of the constituency to whom liberal policies and norms must be justified or explained. A challenge by such integrationist citizens to political liberalism cannot be peremptorily dismissed by suggesting that they are “unreasonable” or fundamentally illiberal and thus beyond the pale of reasonable discussion. The challenge must be taken up and met on its merits.

There is of course one strategy that remains open to Rawls to pre-empt the critique from integrity, namely, to fall back upon a more narrow or restrictive definition of reasonableness that insists upon acceptance of the duty of civility as a condition for entry into the community of reasonable citizens. In my internal critique of political liberalism, I have already considered this strategy and found it to be inconsistent with another key requirement of reasonableness, namely sensitivity to the burdens of judgment (§3.2.2 above). I argued that the duty of civility must be justified within Rawls’s theory to “thinly” reasonable citizens, i.e. citizens who are morally and epistemically responsible and conscientious, willing to offer and abide by fair terms of cooperation, and so on. In my critique of the duty of civility from integrity, I take this conclusion as established.

Finally, even if we accept that Rawls’s “thick” conception of reasonableness (which includes acceptance of the duty of civility) is internally coherent and consistent with Rawls’s other theoretical commitments, my critique from integrity is an external critique and is thus not committed to abiding by Rawls’s definition of reasonableness. Let me state clearly that irrespective of the merits of my internal critique, I believe there are compelling grounds for opting for the thin account of reasonableness over the thick account: The duty of civility depends on prior assumptions that are the subject of a lively
debate in moral philosophy, regarding the relative priority of consent vis-à-vis other moral desiderata such as human well-being, fidelity to religious truth, and the “moral ecology” of society (§3.2.3 above). The only way the thick interpretation of reasonableness can be sustained is by dismissing all thinkers who weight the principle of reciprocity differently to Rawls as unreasonable. But this seems more an arbitrary and self-serving stipulation than an expression of rationality and fairness.

To sum up my interpretation of reasonableness, I have argued that there are compelling grounds, both internal to political liberalism and independent of it, for working with the “thin” definition of reasonableness embedded in Rawls’s account of the burdens of judgment. On this definition, a reasonable person is a sufficiently competent reasoner (the epistemic or rational component) and sufficiently responsive to the demands of justice or the interests of others (the moral component) to be a free and equal citizen in a liberal democratic society.

§3.3.3 The Likely Impact of the Duty of Civility on Reasonable Integrationist Citizens

Now, consider the likely impact of Rawls’s duty of civility upon reasonable integrationist citizens: the duty of civility requires them to (a) identify the political context in which the duty is relevant; (b) distinguish within their own worldview between reasons that are “political” or endorsable by any reasonable person and reasons that are “comprehensive” or not reasonably endorsable\(^\text{190}\); and (c) confine their arguments about

\(^{190}\) Of course, a reason may form part of a comprehensive view and still be reasonably endorsable. However, we may stipulate for convenience that a “comprehensive” reason is a reason that is distinctive of a particular comprehensive view and not a part of public reason.
fundamental law and justice to “political” rather than “comprehensive” or nonpublic, reasons. Now, assume for the sake of argument that integrationist citizens are capable of identifying the cases where the duty of civility is relevant (a); and are capable of recognizing the difference between “political” and “comprehensive” or nonpublic, reasons (b). May an integrationist citizen, without giving up on her integrationist aspirations, consistently confine her arguments about fundamental law and justice to “political” rather than “comprehensive” reasons?

Assuming that integrity requires a certain unity or coherence across our value commitments, and that this unity penetrates every aspect of one’s life, the only way an integrationist citizen could reconcile herself with the duty of civility would be by viewing the primacy of political reasons as flowing from, or at least consistent with, her deepest and most central values. She must find a rationale within her belief system, a rationale that is not at odds with her central commitments, for systematically privileging those reasons within her weltanschauung that are “political” over those reasons that are “personal,” or that presuppose her particular “comprehensive” view, when engaged in debate about fundamental matters of justice. What sort of rationale could that be? Before considering a few examples of integrationist worldviews and how they might view the duty of civility, I propose to set out the most likely rationale for accepting the duty, and I will suggest it is so closely tied to a comprehensively liberal worldview that it is unlikely to be persuasive to citizens who do not share that liberal worldview, even if they are otherwise broadly sympathetic to many traditional liberal democratic values such as liberty, equality, rule of law, due process, freedom of speech, and so on.
Rawls himself acknowledges that at first sight, there is something paradoxical or puzzling about withholding certain claims about reality when deliberating about policies and laws. Why not just bring forward the whole truth, such as it appears to me, when attempting to determine the wisest policy or law? Upon reflection, or so Rawls wants to argue, we can acknowledge that public reason is much like the reason of judges and juries: just as a court must rule out certain evidence if it has not been obtained according to the proper rules—even if the evidence is in fact dispositive of the truth of the matter—a citizen must, in a similar way, screen out certain evidence or reasons for action when debating policies and laws, in this case evidence or reasons that do not meet some sort of publicity test. Otherwise, there is no public, shared basis for the judgment that the legal and political system is morally-grounded and legitimate rather than at the service of this or that person’s peculiar comprehensive views, that have, at best, highly questionable epistemic standing for those who do not share them.

At first glance, this suspicion of justifications for law and policy that fail the publicity test (for now, take that to mean they could not be reasonably endorsed by all reasonable citizens) seems plausible enough. It seems reasonable to suggest that any acceptable rationale for a publicly-binding rule should have some public credibility or epistemic purchase that extends beyond the idiosyncratic worldview of its proponent(s). However, upon further inspection, we find that the duty of civility requires more radical

---

191 The “paradox” of public reason is captured in the question, “why should citizens in discussing and voting on the most fundamental political questions honor the limits of public reason? How can it be either reasonable or rational, when basic matters are at stake, for citizens to appeal only to a public conception of justice and not to the whole truth as they see it? Surely, the most fundamental questions should be settled by appealing to the most important truths, yet these may far transcend public reason!” (PL, 216).

192 For Rawls’s analogy between public reason and legal rules of evidence, see PL, 218-219. For a critique of this analogy, see §4.6, objection 3.
commitments about the nature of political justification and above all, the weight to be attached to freedom from coercion vis-à-vis other values. First of all, the duty of civility is not just an affirmation of the need for some distinction between public and nonpublic reasons, and a privileging of the former over the latter (although some comments by Rawls may seem to imply this); rather, the duty of civility presupposes one peculiar and disputable conception of publicity, embodied in the reciprocity test we have already discussed. According to the duty of civility, reasons for law must be roughly accessible and politically acceptable to any person, whatever his lifestyle, character traits, or philosophy of life, provided he is reasonable or epistemically responsible, cooperative, and amenable in general to political compromise under conditions in which his good will is reciprocated by others. In other words, political coercion must be justifiable to all, not just in the weak sense of grounded in cogent reasons that bear on the common good and the interests of all, but in the much stronger sense that any reasonable person,

193 For example, the following passage suggests that all Rawls is requiring of citizens is that they be prepared to state and defend some minimally plausible criterion of publicity: “The point of the ideal of public reason is that citizens are to conduct their fundamental discussions within the framework of what each regards as a political conception of justice based on values that the others can reasonably be expected to endorse and each is, in good faith, prepared to defend that conception so understood. This means that each of us must have, and be ready to explain, a criterion of what principles and guidelines we think other citizens (who are also free and equal) may reasonably be expected to endorse along with us. We must have some test we are ready to state as to when this condition is met…Of course, we may find that actually others fail to endorse the principles and guidelines our criterion selects. That is to be expected. The idea is that we must have such a criterion and this alone already imposes very considerable discipline on public discussion” (PL, 226-227). But if we are to take this to be Rawls’s considered view of the matter, then much of his own effort to develop a criterion of publicity and explicate the duty of civility based on that criterion seems redundant. Either Rawls’s developed view of publicity is purely suggestive, in which case it has little or no normative bite, or it is normative, in which case it is proposed at the correct and binding criterion. Rawls cannot have it both ways.

194 Even if we accept Rawls’s criterion of publicity, there can be plenty of reasonable disagreement about how to interpret it, since its interpretation is tied to complex moral notions like that of the reasonable person, reasonable comprehensive doctrines, and so on. This is demonstrated in Michael Baur, "On Actualizing Public Reason," Fordham Law Review LLXII, no. 5 (2004).

195 Notice that I am assuming a thinner sense of reasonable that does not imply acceptance of the duty of civility, in accordance with my analysis of Rawls’s account of reasonableness at §3.2.2 and 3.3.2 above.
however situated, would effectively possess a *veto* over political coercion in the event that he could not (reasonably) endorse it, given his current (reasonable) moral and epistemic commitments.

It emerges then that as soon as we acknowledge the duty of civility as a duty, we are already attaching a very high importance to the ability of each and every reasonable citizen to accept the reasons at the bottom of any coercive law which affects them. We are asked not merely to *attempt* to secure the agreement of reasonable persons to the laws which are binding upon them; but to secure that agreement or, failing to do so, give up on the law and effectively give up on the purposes it was meant to serve. Only a comprehensive view such as neo-Kantian contractualism, which affirms the moral primacy of *reasonable consent* over and above other moral values such as ethical well-being, the common good, religious salvation, etc., could give such a central and decisive role to reasonable consensus in the justification of law and policy. This point should become clearer as we consider a few examples of integrationist citizens and how they fare in a regime structured by public reason.

As MacIntyre points out, moral principles are not fully grasped until we have envisaged how, if at all, they can be implemented within a real society. In short, he argues for the relevance of sociological observation and thought to moral and ethical reflection. This is a point that Rawls is not as attentive to as he might be. For example, he does not explore in any adequate way the sociological implications of public reason. But

---

196 See Alasdair MacIntyre, *After Virtue: A Study in Moral Theory* (London: Duckworth, 1981), 23: “we have not yet fully understood the claims of any moral philosophy until we have spelled out what its social embodiment would be. Some moral philosophers in the past, perhaps most, have understood this spelling out as itself one part of the task of moral philosophy. So, it scarcely needs to be said, Plato and Aristotle, so indeed also Hume and Adam Smith; but at least since Moore the dominant narrow conception of moral philosophy has ensured that the moral philosophers could ignore this task…”
we can only fully grasp the *normative* significance of public reason by understanding how it might *work* in an actual society something like the liberal democracies of Europe and North-America.\(^{197}\) Our discussion so far has focused on aspects of the duty of civility, at the theoretical level, that suggest it may be exceedingly hard to swallow for many integrationist citizens. I now want to imagine a society in which the duty of civility has become effectively rooted in its public culture, in order to explore its concrete implications for integrationist citizens.

Recall that according to the duty of civility, citizens have the duty “to be able to explain to one another on...fundamental questions [basic justice and constitutional essentials] how the principles and policies they advocate and vote for can be supported by the political values of public reason.”\(^{198}\) Now, let us imagine a liberal democratic culture, much like our own (say, the United States) except approximating what Rawls calls the “well-ordered” society\(^ {199}\), in which Rawls’s ideal of public reason and its associated duty of civility become a dominant moral force in the public culture. My analysis will focus on the Rawlsian duty of civility and its implications for political speech.

---

\(^{197}\) My discussion presupposes a north-Atlantic constitutional democracy not because I believe liberalism is irrelevant to other societies, but because this is what Rawls himself presupposes. For an attempt to extend the most basic principles of political liberalism to non-democratic cultures, see John Rawls, *The Law of Peoples* (Cambridge, Massachusetts and London, England: Harvard University Press, 1999).

\(^{198}\) PL, 217.

\(^{199}\) The well-ordered society is one in which (a) “everyone accepts, and knows that everyone else accepts, the very same political conception of justice (and so the same principles of political justice),” (b) “society’s basic structure—that is, its main political and social institutions and the way they hang together as one system of cooperation—is publicly known, or with good reason believed, to satisfy those principles of justice,” and (c) citizens have a normally effective sense of justice, that is, one that enables them to understand and apply the publicly recognized principles of justice, and for the most part to act accordingly as their position in society, with its duties and obligations, requires” [*Rawls, Justice as Fairness: A Restatement*, 8-9]. The reason I say “approximating” rather than instantiating the well-ordered society, is that the notion of universal acceptance of a single political conception is clearly an idealization that one cannot envisage occurring given the human condition and what Rawls calls the “burdens of judgment.”
In such a society, the Rawlsian ideal of free and equal citizenship is accepted by a majority of citizens along with its associated norms and expectations concerning good citizenship. What I would like to highlight here is the so-called “duty of civility,” the moral (not legal) duty that citizens have to restrict their public advocacy of legislation and public policies, at least when these involve questions of basic justice or constitutional essentials, to arguments that can be supported “in due course” by “public reasons, or the political values covered by the political conception of justice (or one of a suitable family of such).” A culture in which public reason is well-established is a culture in which citizens expect from each other political speech that complies with the duty of civility. In other words, the duty of civility is a powerful and pervasive social norm, and failure to adhere to it normally occasions the judgment that one is a less than fully responsible, thoughtful, cooperative, or reasonable citizen. In short, the social sanctions attaching to noncompliance—the moral disapprobation of one’s peers—are powerful enough to place considerable external pressure on citizens to conform, irrespective of their own personal views on the matter.

Now, in order to explore the implications of the duty of civility for integrationist citizens, I will speculate about its effects on the lives of four citizen “types” whom I have

---

200 Ibid., 90. Of course, this does not mean that one is morally precluded from advocating laws based on nonpublic reasons, but it does mean that one must find public reasons “in due course.” It is not clear to me how the so-called “proviso” of “in due course” significantly softens the duty of civility. After all, surely a citizen cannot advance nonpublic reasons unless he or she at least has grounds for believing, in good faith, that comparatively persuasive public reasons will be forthcoming sooner rather than later? This is not the place to argue the merits of the duty of civility, merely to point out that it still has significant bite even after the proviso has been added.

200 Note that three of the four types, the Kantian, the Millian and the Catholic, are close to the citizens deemed by Rawls potential candidates for an “overlapping consensus” on the principles of justice, that is, citizens who could conceivably affirm the public conception of justice from within their own comprehensive doctrines: the Kantian, the utilitarian, and the religious believer. On this point, see PL, 169-170.
invented but who are meant to represent at least some of the views of real citizens.201 There may be many citizens who differ from the types in certain details but nonetheless resemble them sufficiently to be similarly affected by the duty of civility. Notice that the four types in question are all united in accepting some fairly robust ideal of liberal democracy – they all accept, or can be assumed to accept, a liberal constitutional order which guarantees citizens certain rights and liberties, and they are willing to participate in and accept the outcome of democratic procedures, if not absolutely, at least within broad parameters. All four of our citizen types are reasonable in at least two senses: (a) epistemically, by being willing to alter their views in response to evidence, willing to give genuine consideration to the arguments of others, following basic rules of logic and rationality, etc.; and (b) morally, by being willing to offer others and comply with what they judge in good faith to be fair terms of cooperation, and in general being sensitive to the needs and interests of their fellow citizens.202

1. First, there is the Millian liberal who believes that the goal of social and political order is the cultivation of individual talents and the accommodation of rare genius, but above all the improvement of the human condition all round. The Millian will have a progressive view of history, or at least he views the possibility of human progress under rationally constructed liberal institutions with considerable optimism. The Millian liberal believes that the public arena should be receptive to wildly diverse claims and

201 Note that three of the four types, the Kantian, the Millian and the Catholic, are close to the citizens deemed by Rawls potential candidates for an “overlapping consensus” on the principles of justice, that is, citizens who could conceivably affirm the public conception of justice from within their own comprehensive doctrines: the Kantian, the utilitarian, and the religious believer. On this point, see PL, 169-170.

202 This is basically what I call Rawls’s “thin” sense of reasonableness (see §3.3.2 above).
arguments, not for the sake of diversity, but because he believes that only in a robust and free exchange of views and arguments is the truth likely to emerge. Furthermore, he believes that it is only when confronted with the provocative and strange beliefs of one’s fellow citizens that one is forced to reflect in a serious way upon one’s own beliefs rather than blindly following public opinion.203

2. Second, there is the Kantian who believes in certain universal and unchanging norms governing human behavior, most centrally the norm that we should never treat any human being as merely a means but always as an end it himself. Practical reason provides us with a set of categorical as well as hypothetical imperatives which can provide sound moral guidance to us both individually and in our communal decisions. One of the central goals of law, according to the Kantian citizen, is to create a social environment in which moral autonomy can flourish.204

3. Third, there is the natural lawyer, who believes that the medieval tradition of natural law as recently reinvigorated by late twentieth century thinkers, provides the most defensible account of the human good, and can provide genuine guidance to clear thinking on a series of social and moral issues, including questions of basic justice and constitutional essentials. Furthermore, the natural lawyer believes that any attempt to divorce positive law from natural law, or to dispense with the idea of a God before whom we are accountable for our actions, tends to reduce the law to a self-grounding human contract with no deep moral justification. The natural lawyer does not rest her case on revealed theology but argues, as in natural theology, that we can see, at least

203 This picture of the Millian liberal is based roughly on J.S. Mill’s On Liberty, esp. ch. III.

probabilistically, the effects of God’s nature and existence in the universe and in our own natures.\footnote{For a recent defense of a theistic natural law position, see Hittinger, \textit{The First Grace}.}

4. Fourth, there is the Catholic volunteer, who works for a crisis pregnancy center and believes that our policies regarding single and unwanted pregnancies should be modeled on the compassion of Christ in the Gospel. Only by experiencing God’s forgiveness and compassion in our lives, according to this Catholic, can we truly formulate policies and laws that reflect an attitude of care and compassion and understanding. Because of his Christian values, he believes that laws should be passed that promote generous educational and charitable initiatives, not primarily under the State’s supervision, but in collaboration with charitable organizations that are better placed to meet the needs of women.

Notice that only one of the types I have sketched falls under the category, “religious” (at least in the sectarian or confessional sense).\footnote{The position of the natural lawyer may be viewed by some as religious, but it is not religious in the sense of confessional or credal. It relies on natural theology and purports to rest on the evidence of nature rather than revealed theology or scripture.} I want to show that all four of the citizen types I have sketched will often be (socially) penalized for their efforts to live good lives (lives of integrity, if you will) in a culture in which public reason is a dominant social and moral force. In order to do this, I will consider how our four citizen types would fare in their pursuit of integrity in a political culture regulated by the Rawlsian ideal of citizenship and its associated duty of civility:
1. The Millean liberal will not fare well under public reason. First, he will view public opinion and prevailing beliefs with great suspicion, since he is convinced that in a democratic culture, people tend to accept common beliefs blindly rather than developing well-grounded views of their own. In light of the general tendency towards passive acceptance of the status quo, the Millian liberal will probably view his engagement in public discourse as an opportunity to shock his compatriots into re-evaluating their own beliefs. Therefore, rather than simply rehearsing the shared values of his political culture, he will likely feel morally obligated to challenge many of those values, in order to awaken his fellow citizens from their complacent slumber and their intolerance of those who dare to question prevailing notions. A Millian liberal would be unlikely to comply willingly with the duty of civility: on the contrary, he would likely feel it as an onerous burden upon his pursuit of his conception of the good in and through discourse.207

2. Now, what of the Kantian? It seems that at least some Kantians will insist that the deepest foundations of morality can and must be exposed in the course of honest deliberations about political matters. Whether or not everybody shares his Kantian moral convictions, the Kantian is convinced that they are the best available moral argument for a liberal society and he may even believe that watered-down versions of Kantian arguments are quite inadequate and potentially harmful or debilitating for liberal institutions. For a Kantian, the moral truth contained in the dictates of practical reason, though available in principle to everyone, can be obscured by our passions and desires,

207 The harm principle may well justify certain restrictions, whether moral or legal, of free speech, but it is hard to see how the harm principle could be used to justify such a far-reaching (social/moral) restriction on speech as the duty of civility. For a good contemporary defense of a Millian conception of public reason that opposes itself to Rawlsian public reason, see Jeremy Waldron, "Religious Contributions in Public Deliberation," San Diego Law Review 30 (1993).
but this is no reason to desist from propounding that truth and garnishing support for it politically. To expect a Kantian of this sort to desist from arguing for policies unless they can be supported on independent “public” grounds might well condemn him to silence, since he may well believe that Kantian morality provides the only adequate grounds for many morally sound policies. While a Kantian might accept some moderate version of the burdens of judgment as part of an explanation for moral divergence, he would insist that the moral law in its essentials is accessible to any rational person, and that insofar as people fail to see it, this is on account of some obstacle not reducible to the burdens of judgment, such as corrupt upbringing or selfishness. But the law, though it must take into account the baser and more contingent elements in human nature, must nonetheless persevere in promoting and protecting the moral law with or without the support of all.

3. Now consider the natural lawyer. Although she believes that natural law is the only sound basis for a healthy political society, she acknowledges that it is a notion that is widely discredited and she believes that our culture is for various reasons unreceptive to it. Nonetheless, she thinks it is important to continue to argue for and from natural law principles even in the face of deep disagreement and controversy, since this is the only way that they have a chance of eventually becoming “mainstream.” The natural law embodies the truth about politics as the natural lawyer sees it, or at least it provides essential tools for getting at that truth. As such, she feels bound by her own duty to promote the real (as opposed to merely apparent) good in society, to advance natural law

---

208 Of course, the notion of public reason itself finds one of its earliest prototypes in Kant. However, in Kant it was a concept of pure practical reason and as such was not conditioned by the intuitions or prevailing views of a culture. See, for example, appendix 2 to “To Perpetual Peace A Sketch” [1795], “On the Agreement between Politics and Morality under the Transcendental Concept of Public Right,” in Immanuel Kant, Perpetual Peace and Other Essays, trans. Ted Humphrey (Cambridge and Indianapolis: Hackett Publishing Company, 1983), 135-139 [AK 8:381-386].
arguments for policies and laws. Furthermore, she does not believe there are more “common-sensical,” “public” reasons that can yield the same conclusions as natural law arguments. It is hard to see how anything this philosophically sophisticated and contentious could be considered part of the “common sense” of society or how its basic assumptions could be considered already “implicit in the political culture” of her society. This would seem to suggest that many natural law arguments would fall foul of the duty of civility. Thus, to comply with that duty would condemn the natural lawyer to political speech that she considers superficial and of very limited moral value, since she would have to forego her commitment to what she views as an essential step towards the long-term improvement of society, namely, a revival of natural law reasoning in matters of law and public policy.

4. Finally, consider the Catholic volunteer. It is possible that in certain situations, he might think it more effective to campaign for his favored government policies towards pregnancy on non-religious grounds. However, he will also likely believe that it is important that such policies be formulated in the spirit of Christ’s infinite compassion for the downtrodden, and he may believe that “public” reasons alone do not provide fully adequate grounds for the policies he favors. He may believe that Christ provides the model of compassion that is most appropriate for policymakers in this area. He would

---

209 Of course, if we allowed a great deal of latitude to this term, “implicit” and its interpretation, then a wide variety of comprehensive doctrines could reasonably be viewed as “implicit” in our public culture. But then, Rawls’s distinction between political and “comprehensive” doctrines would be otiose.

210 That natural law arguments for statutes or policies on questions of basic justice fall outside public reason and are thus politically illegitimate is clear from Rawls’s comment that “in considering whether to make homosexual relations between citizens criminal offenses, the question is not whether those relations are precluded by a worthy idea of full human good as characterized by a sound philosophical and nonreligious view, nor whether those of religious faith regard it as sin, but primarily whether legislative statutes forbidding those relations infringe the civil rights of free and equal democratic citizens” (Rawls, "The Idea of Public Reason Revisited," 588). One of the basic premises of classical natural law theory is that legal and moral norms are ultimately grounded in “a worthy idea of full human good.”
likely resent the suggestion that he is failing to be a good citizen by arguing from Christ’s compassion for more generous and thoughtful action by legislators on behalf of women experiencing crisis pregnancies. His consistent adherence to the duty of civility might well preclude him from fully living out his specifically Christian commitment to the poor and the marginalized in his public advocacy.

In each of the cases discussed above, public reason effectively places an informal, social restriction upon speech-acts that are viewed by citizens as essential constituents of their moral integrity, that is, such speech-acts as are thought by citizens to be consequential upon, or required by, one or more of their most fundamental moral beliefs. Even granting that public reason does not prevent citizens from engaging in similar speech-acts in the “background culture”\(^\text{211}\) or in the context of voluntary associations in which they participate, the citizens we have considered will view this selective constraint not as a constraint required by practical reason itself but as a source of moral dis-integrity, i.e., as a temptation to betray one’s moral convictions in the public arena in order to avoid the social sanctions attached to more integrated behavior. Though social restrictions do not function identically to legal restrictions, they do function in closely analogous ways: a certain penalty falls upon the offender—in this case, social and moral disapprobation—and that penalty is not an “act of God” but is imposed, whether singly or jointly, by other human agents.

\(^{211}\) Rawls stipulates that the “background culture includes…the culture of churches and associations of all kinds, and institutions of learning at all levels, especially universities and professional schools, scientific and other societies. In addition, the nonpublic political culture mediates between the public political culture and the background culture. This comprises media—properly so-named—of all kinds: newspapers, reviews and magazines, television and radio, and much else” (Ibid., 576, fn. 13).
Of course, Rawls might respond that his ideal of citizenship is not a legally enforceable ideal, but a voluntary one, that is, an ideal to be adopted by citizens upon due reflection, should they judge it to be compatible with their other commitments: in short, citizens are to embrace the ideal of citizenship voluntarily, provided they judge it to be compatible with their integrity. This insistence on the voluntariness of the ideal at first sight appears to soften if not entirely blunt the tension between the duty of civility and the pursuit of integrity. If citizens are free to accept or reject the duty, then there is no general incompatibility between integrity and the recognition by some of the duty of civility. After all, Rawls does not insist that his ideal be embraced willy-nilly by each and every citizen. He at no point claims that the political conception of justice and its attendant norms are absolutely and in all cases overriding with respect to comprehensive doctrines. On the contrary, the ultimate judgment call must be made by individual citizens in good conscience.

Now, if the story were that simple, the duty of civility would indeed appear to pose a much weaker threat to integrity than it did at first blush. However, the story is not so simple. The duty of civility is not just a duty that some citizens will feel compelled in conscience to opt out of, but one that, should it gain a foothold in the political culture, will lead to widespread moral disapproval of the behavior of citizens who fail to comply. These social sanctions, though not identical to legal sanctions, are likely to operate in similar ways: people will anticipate the social sanction of disapproval should they violate

---

212 For example, “the basis of citizens’ respect for the limits of public reason,” which limits are clearly a central component of Rawls’s ideal of citizenship, is “a balance of reasons as seen within each citizen’s comprehensive doctrine, and not a compromise compelled by circumstances” (PL, 169). And elsewhere, Rawls stresses that “citizens individually decide for themselves in what way the public political conception all affirm [which includes the duty of civility, we might add] is related to their own more comprehensive views” (PL, 38). In addition, Rawls gives a very full-blooded defense of freedom of speech (PL, 340-348) which, at least in its legal dimensions, is not overridden by the (moral, not legal) duty of civility.
the duty of civility, and this expected cost of violation will often make compliance considerably more attractive than non-compliance. Thus, the supposed “voluntariness” of the duty of civility is belied by the most obvious sociological mechanisms through which the duty can become effective. Indeed, it is hard to imagine the duty of civility functioning in any other way in a society that is broadly sympathetic to political liberalism.

§3.4 Implications of the Duty for Political Stability

In addition to failing Rawls’s own reciprocity test and severely burdening the personal integrity of reasonable citizens, the duty of civility, one of the centerpieces of public reason, seriously jeopardizes one of the three central goals of political liberalism, namely, to win a sufficient degree of support among citizens to be stable or provide a reliable basis for public order.\textsuperscript{213} It is not just that the duty of civility is deeply problematic from the point of view of citizens who reject the fundamental values of constitutional democracy, but it is deeply problematic from the perspective of many citizens who accept, to a large extent, and are willing to work within, a liberal democratic political order. I have already shown how four citizen of a roughly integrationist persuasion whose reasonableness and cooperativeness is hard to question, can only comply with the duty of civility at the cost of abandoning some of their central ethical commitments.\textsuperscript{214}

\textsuperscript{213} For the three central goals of political liberalism, see §2.2 above.

\textsuperscript{214} Given the contestable epistemological and normative claims underpinning the duty of civility, there may also be many non-integrationist citizens who cannot comply with the duty without betraying some central ethical commitments. But I do not wish to press this point here.
This presents integrationist citizens, who represent at least a significant minority in the United States which Rawls’s theory is primarily intended for, with a serious dilemma: should they choose to engage in political advocacy, at least with respect to matters of basic justice or constitutional essentials, they must either remain true to their central ethical commitments and flout the duty of civility, in which case they will be considered to be breaking the rules of the game, and falling short as citizens; or they can accept the rules of the game, conform to the Rawlsian ideal of citizenship, and win the approval and acceptance of their peers, but find that in so doing, they have betrayed central ethical commitments and perhaps done so under social duress.

Where integrationist citizens refuse to comply with the duty of civility, they will suffer general disapprobation and the \textit{bona fides} of their contributions will come under question, independently of the intrinsic merit of their arguments. They will be disadvantaged in politics on account of their integrationist commitments and the seriousness of their commitment to liberal democratic values will be impugned. They may continue to believe themselves to be acting in good faith and to be living up to the true requirements of citizenship, but they will likely feel alienated from the wider political process and unfairly slighted by the dominant ideal of citizenship.\footnote{A few years ago, a European politician (Rocco Buttiglione) lost his candidacy for a leadership position in parliament for admitting that he believed as a Catholic that homosexuality was disordered, even though he unambiguously affirmed his commitment to non-discrimination against homosexuals. Interestingly (especially given the outcome), he was in fact complying with the duty of civility since he was basing his political decisions on public reasons (non-discrimination). See Rocco Buttiglione’s presentation at the sixth \textit{Congress on Catholics and Public Life} in Madrid, Spain, on Saturday, November 20, 2004: “The Buttiglione case as seen by Buttiglione,” \url{http://www.secondspring.co.uk/articles/buttiglione.htm}.} Where integrationist citizens comply with the duty, they may feel accepted by their fellow citizens and somewhat acceptable to the wider political culture, but alienated from some central ethical commitments of their own or of their community’s. They may feel
unwilling to own up to their most important ethical beliefs or commitments in a public culture that will penalize them socially for insisting on pursuing value-integrity in their political speech. Compliance will not be a matter of wholehearted acceptance of the duty of civility, but reluctant acceptance of the duty as the inevitable price of social and political success and esteem. One of the residual effects of this reluctant or half-hearted compliance will likely be a sense of personal betrayal or inauthenticity and a substantial degree of resentment towards a cultural and political environment that makes this inauthenticity the prerequisite for political acceptance or correctness.

The two scenarios we have explored so far, those of political participation marked by compliance and non-compliance with the duty of civility, are both likely to generate a substantial degree of resentment among integrationist citizens towards the wider political culture, and a general sense of alienation from the political process, which is so unfriendly to their ethical outlooks. Whether they have chosen to comply with the duty of civility or not, integrationist citizens who engage the political process will feel disadvantaged and marginalized by the process, and may frequently assess the legitimacy of the outcome with suspicion or at least with significant reservations.

Of course, there will also be those integrationist citizens for whom the price of political participation, whether in the role of rebel or in the role of reluctant conformist, is too high. They may feel the strain of opposing commitments, or the stress of societal disapproval and marginalization, is not worth the benefits of participation. In that case, they will simply disengage from the political process, or only enter it occasionally, in skirmish-like fashion. Their allegiance to the wider polity will be weakened as they form
their own communities and adopt a stance of stoic resistance or detachment towards the wider political culture.

To sum up, in a political regime structured by public reason, where the duty of civility is widely recognized to be binding upon citizens, integrationist citizens will find themselves alienated and possibly resentful towards the political process whether they (a) engage it on its own terms; or (b) engage it as rebels, playing outside the rules of the game; or (c) withdraw from the process but still find their lives are deeply affected by its outcomes (primarily laws). Assuming only that integrationist citizens make up a significant minority of the population of the United States, the country for which Rawls’s ideal of citizenship is primarily intended, Rawls’s ideal of citizenship, or at least one central component of it, public reason, seriously jeopardizes one of the practical conditions political liberalism is meant to promote, namely “an enduring and secure democratic regime…not divided into contending doctrinal confessions and hostile social classes,…willingly and freely supported by at least a substantial majority of its politically active citizens.”

§3.5 Rethinking Reciprocity

I have examined Rawls’s ideal of citizenship specifically as it affects citizens’ discussions about matters of basic justice and constitutional essentials, and found it wanting on at least two counts: first, it fails to accommodate, indeed significantly burdens, reasonable and generally cooperative citizens’ efforts to pursue the good in an integrated way. Second, Rawls’s ideal of citizenship is deeply incoherent even on its own

\[216\] PL, 38.
terms, since there is no reason to believe that the duty of civility, given its underlying reciprocity requirement, could be accepted by any “reasonable person” as a constraint on political justification, except by definitional sleight of hand or by arbitrary stipulation. Given the controversial philosophical grounds of the duty of civility, to deem dissenters from this requirement “unreasonable” would be effectively to require acceptance of a particular comprehensive doctrine or controversial worldview (however partial it may be) as a condition for entry into the community of “reasonable,” cooperative citizens. But this is precisely what the duty of civility was designed to avoid. Thus, as a political norm, the duty fails Rawls’s own test of legitimacy.

However, even if a Rawlsian is not convinced that the reciprocity requirement is self-defeating in the way I have suggested, the consistently agonistic experience of practical politics and political philosophy alike should be enough to convince us that any interesting, informative, and consequential vision of justice, at least under modern conditions of pluralism, will fail the test of reciprocity. First, the very “fact of reasonable pluralism” that is so central to Rawls’s turn from philosophical to political liberalism gives us reason to be skeptical of any effort to formulate a normative foundation for politics that somehow rises above the fray of reasonable disagreement. Second, all efforts that I am aware of to articulate informative and usable normative principles to govern political practices, whether aiming to be simply true (say, Raz’s autonomy-based “perfectionist” theory\(^ {217} \)) or aiming to be acceptable to any suitably motivated agent (say, Rawls’s political liberalism), have met with radical disagreement and controversy, not only outside the liberal tradition (say, the concerns of neo-Aristotelians and Thomists),

but from within the liberal tradition itself (consider Dworkin’s rejection of what he calls Rawls’s “strategy of discontinuity”\textsuperscript{218} and Galston’s explicit repudiation of Rawls’s attempt to be “political, not metaphysical.”\textsuperscript{219}). The consistent failure of political theory to avoid this sort of deep-seated controversy in the past, even among those with an avowedly liberal worldview, suggests that any hope of doing so in the future is deeply misguided.

In short, what Rawls calls “reasonable disagreement” is in fact more pervasive and far-reaching than Rawls himself is willing to admit. If I am right, and no interesting justification of liberal democratic practices is likely to pass the test of reciprocity as formulated by Rawls, then we have three options: on the one hand, we could accept that political theory and political regimes are inevitably illegitimate (on the reciprocity test) and consistently violative of citizens’ autonomy. In that case, the quest for political legitimacy appears hopelessly futile and we must simply engage in some form of damage limitation (say, reduction of suffering, or promotion of goods like health and education) within a fundamentally illegitimate web of ideas and practices. A second option is to make the test a matter of degrees rather than of some on/off property. On this view, the test of reciprocity, though it must inevitably fail in its strongest form, could be met to some limited extent, e.g. we could achieve \textit{more} rather than \textit{less} reasonable agreement, we could achieve agreement on \textit{more} rather than \textit{less} of our political principles and practices.\textsuperscript{220} However, given the extent of disagreement that political liberalism has met,

\textsuperscript{218} See Dworkin, “Foundations of Liberal Equality,”
\textsuperscript{219} See Galston, \textit{Liberal Pluralism} .
\textsuperscript{220} In practice, of course, the test of reciprocity is never applied in an absolute way, such that \textit{every single reasonable citizen} must be able to affirm the principles that structure a regime; as Klosko suggests
both within and without the liberal tradition, the test of reciprocity seems likely to fail in such a large proportion of cases that maximizing reciprocity looks like a euphemism for resigning ourselves to failure.

A third strategy we might adopt in the face of the consistent failure of our political practices and theorizing to meet the Rawlsian test of reciprocity is to subject the test itself and its philosophical presuppositions to a fundamental re-examination, and search for some other notion of reciprocity that might have some chance of (morally) regulating the actual political practices of our world. This is the strategy I favor, and it involves a re-examination not only of reciprocity, but of closely related ideas such as political autonomy, respect, and reasonableness. Although this is a much larger project than I can undertake in this dissertation, I aim to at least begin such a re-examination in chapters 5 and 6, prompted not only by the contradictions internal to political liberalism, but also by the need to develop an ideal of citizenship that accommodates rather than punishes citizens’ efforts to live lives of integrity. But first, we must think through at a deeper level what the notion of integrity amounts to, and how it can overcome some powerful liberal objections (chapter 4), since the overarching goal of this project is to make room for a politics of integrity.

221 By now, the extent of deep disagreement among academics on the merits of political liberalism as an approach to politics is so great that I will only mention a few representative dissenters who are widely considered to be operating within the liberal tradition and whose preference for liberal democracy cannot be disputed: Dworkin, “Foundations of Liberal Equality,”; Galston, *Liberal Pluralism*; Raz, "Facing Diversity,"
CHAPTER 4: THE CONCEPT OF INTEGRITY

The Rawlsian ideal of citizenship has been criticized from different quarters for failing to accommodate a citizen’s legitimate aspiration to make sense of his life as a whole in terms of his deepest ethical convictions. This criticism has been leveled against political liberalism both from the perspective of a more “ethical” or “comprehensive” liberalism\textsuperscript{222} and from the perspective of a religious believer who owes her primary allegiance to God and His Will (however it may be revealed) rather than to any “political” conception of justice.\textsuperscript{223} As I have argued in chapter 3, Rawls subordinates integrity to the principle of reciprocity, and the value of integrity as a dimension of the moral life can occupy at best a derivative and heavily circumscribed role within political liberalism. For Rawls, this is just a function of the normative demands of political life, and as such, it is a problem shared by any viable conception of justice; while for his opponents, Rawls’s attack on integrity is much more aggressive and far-reaching than it needs to be: they believe, or at least hope, that a conception of justice can be considerably more accommodating (if not completely accommodating) of integrity than Rawls’s.

\textsuperscript{222} See, for example, Dworkin, “Foundations of Liberal Equality,” which seeks to establish greater “continuity” between our ethical and political commitments than Rawls will allow.

I agree with Rawls’s critics that political liberalism constitutes an aggressive and unnecessary attack on the integrity of citizens. I, too, am hopeful that a politics that is considerably more accommodating of integrity is possible. However, at least up to this point, critics of political liberalism have been somewhat elusive about the concept of integrity itself, and its role in the moral life. Their failure to tackle the concept head-on leaves the case against the Rawlsian ideal of citizenship, insofar as it is inspired by the value of integrity, on very rocky ground indeed. This is because some of the most interesting and challenging questions about integrity and its role in political life cannot be adequately addressed until political proponents of integrity become more perspicuous about what they mean by integrity.\footnote{By political proponents of integrity, I mean those who advocate the accommodation of integrity by political authority and self-consciously deny the Rawlsian-liberal attempt to cordon off comprehensive doctrines from political conceptions of justice.} This would require paying attention to what moral philosophers have had to say about integrity, building up a concept of integrity informed by that discussion, and exhibiting some of its most important implications for political and social life. Only then, when political proponents of integrity have developed a rich and consequential understanding of integrity, can they begin to answer the sorts of questions a political liberal is likely to raise, such as: why should we care so much about integrity, if it is largely independent of more substantive moral ideals? Isn’t it impractical and even foolish to think that a political regime could accommodate the integrity of all of its citizens, given the range of ideals and values they espouse? Doesn’t a preoccupation with integrity license citizens to disregard political justice whenever their “integrity” comes into conflict with freedom or justice? And if people’s political speech is fully
“integrated” with their deepest values, isn’t this likely to heavily corrode the prospects for constructive political discourse in a morally heterogeneous society?

In response to these rather pointed questions, it is not enough to simply point to the fact that many people do seek to live a highly integrated life. What has to be shown instead is that the aspiration to live an integrated life is not unreasonable, and is compatible with a political order that is both adequately stable and adequately sensitive to certain fundamental values associated with liberal democracy such as personal freedom, political liberty, and equality before the law. In order to respond to this challenge (what I call the “challenge of citizenship”), integrationist critics of political liberalism need to not only develop a positive concept of integrity that is responsive to the challenge of citizenship: they also need to develop an alternative ideal of citizenship that (a) is morally persuasive and plausible under modern political conditions, and (b) accommodates the personal integrity of citizens. I defer the second task—that of developing an alternative ideal of citizenship—to the last chapter. Meanwhile, in this chapter I aim to elaborate a concept of integrity that neither issues in political anarchy nor in a heavily paternalistic, or even totalitarian, regime (after all, what if people attempt to impose their deepest moral values upon the polity at large, attempting to order everyone’s life according to their peculiar vision of the good? What if an integrated moral life gives rise to an equally integrated public morality?).

It is primarily with the challenge of citizenship in view that I embark upon the difficult task of spelling out a richer account of integrity than those offered by its political proponents thus far. I will engage ethical discussions of integrity with three principal goals in mind: (a) to make some modest contribution to our understanding of this concept
and its role in the moral life, which is a matter of independent interest to moral philosophers; (b) to buttress the case against the Rawlsian ideal of citizenship, by making a preliminary showing that the unreserved pursuit of integrity is *politically and morally reasonable*; and (c) provide normative parameters for an alternative, aretaic ideal of citizenship. It will be my contention that no viable ideal of citizenship can afford to reject or de-legitimate the pursuit of integrity as I understand the term.

I will preface my treatment of integrity with (§4.1) a rough characterization of the core concept of integrity and its application to moral agents; and (§4.2) a brief description of some essential features of the moral agent to whom we may intelligibly ascribe personal integrity. (§4.3) I will then draw a rough taxonomy of the different approaches to integrity to be found in the extant Anglo-American philosophical literature, and (§4.4) argue that most available accounts of integrity are beset by serious difficulties, in particular a lack of conceptual rigor or unity. Finally, (§4.5) I will set about refining my initial characterization of integrity, and begin to develop a conception of integrity that aims to surpass most existing accounts by explaining the connection between the *formal* and *substantive* dimensions of this virtue. I will enhance this exposition by (§4.6) articulating and responding to four key objections that might be leveled against the virtue of integrity by a liberal theorist such as Rawls.

---

225 This case remains incomplete until I have laid out a virtue ethical model of citizenship that aims to accommodate integrity as I describe it in this chapter. I do not undertake that task until the final chapter.

226 Since this term, “aretaic” is not used by everyone, I should mention that by “aretaic ideal,” I simply mean an ideal rooted in or constituted by an account of the virtues (rather than, say, a set of principles).
§4.1 A Preliminary Exploration of Integrity

The word “integrity” has many different albeit related meanings, depending on the object described and the field of inquiry. For example, we can speak of the “integrity” of an institution (“this decision impugns the integrity of our highest court”), a bridge (“this bridge is badly designed – it lacks structural integrity”), a procedure (“we must ensure the integrity of the electoral process”), or a person (“when he told the truth at such a high personal cost, he displayed admirable integrity”). While applications in one context may provide some insights about applications in another, the main focus of this inquiry is the concept of integrity as a property of human beings or persons. Furthermore, I am not interested in integrity as a physiological or biological property but as a psychological and moral attribute, a property we ascribe to human persons to characterize in some way the quality of their moral agency, viz., their capacity to act not just unreflectively, like an animal which responds instinctively to its environment, but reflexively, as a creature that can not only deliberate about what to do, but can become aware of its own choices and desires, endorsing or disapproving of them, and can view its life as a discrete whole rather than merely as a sequence of discrete memories.

Granting, then, that integrity is a property of persons viewed as moral agents, a property that tells us something about the quality of their moral agency, what sort of

---

227 Our ability to step back from and evaluate our desires entails the capacity to have second-order desires (desires about desires). Practical evaluations, though not identical to desires, do seem to imply the presence of a desire (sometimes a retrospective wish) that such-and-such an action or desire come about, or that such-and-such an action or desire cease. Desires, on the other hand, imply evaluations whether express or inchoate, insofar as one cannot desire an object without implicitly or explicitly judging it to have some value. The distinction between first- and second-order desires originates in a well-known essay, Harry Frankfurt, "Freedom of the Will and the Concept of a Person," *The Journal of Philosophy* 68, no. 1 (1971). Taylor builds on the distinction, and introduces a new one, that between "strong evaluative" and "weak evaluative" desires: Charles Taylor, “What Is Human Agency?,” in *The Self: Psychological and Philosophical Issues*, ed. Theodore Mischel (New Jersey: Rowman and Littlefield, 1977).
property might this be? A good place to start is the Oxford English Dictionary: The word “integrity” is derived from the Latin word, “integritas,” meaning “wholeness, entireness, completeness, integrity, chastity, purity,” and the corresponding noun, “integer,” meaning “whole.” Notice that there are two senses of integrity, a more generic sense, “wholeness, undividedness,” and a more specialized moral sense, “the condition of being uncorrupted, innocence, the character of uncorrupted virtue, honesty, uprightness, fair-dealing, etc.” The connection between the literal and the moral sense of integrity is not self-evident. However, since it is the critique of political liberalism from integrity that this inquiry is attempting to support, and since that critique consistently appeals to notions of consistency and wholeness, it makes sense for us to focus on the concept of integrity as wholeness, and inquire what connection if any, that sense of integrity has to the moral sense of the term.

But there are also independent reasons for beginning our inquiry into integrity with the concept of wholeness rather than, say, honesty or uprightness. First, if we could apply the generic, literal meaning of integrity to the moral life, we would have a more

---

228 Examples given by OED are: “1533 MORE Answ. Poysoned Bk. Wks. 1095/1 Not yᵉ sacrifice nor oblacion, whyche to the integritie thereof requyreth both the formes. a1677 HALE Prim. Orig. Man. I. iii. 93 Are there not among men some that want the integrity of their Limbs? 1756-82 J. WARTON Ess. Pope (ed. 4) I. III. 101 The poem before us is by no means destitute of a just integrity, and a lucid order. 1833 LYELL Princ. Geol. III. 192 The integrity of the cones..shows that the country has not been agitated by violent earthquakes. 1838 SIR W. HAMILTON Logic xxiv. (1866) II. 2 Method, considered in its integrity, consists of two processes,—Analysis and Synthesis. 1868 G. DUFF Pol. Surv. 44 To work for the integrity of the Ottoman Empire, and at the same time to promote its dismemberment, is to stultify one's-self. 1870 F. R. WILSON Ch. Lindisfarne 68 The walls were standing..though not in their integrity.”

229 Examples given by OED are: “1548 HALL Chron., Hen. VI 185b, So much estemed..for his liberalitie, clemencie, integritie, and corage. 1599 Life More in Wordsw. Eccl. Biog. (1853) II. 157 That he might reserve the integritie of a good conscience. 1611 BIBLE Prov. xix. 1 Better is the poore that walketh in his integrity, then he that is peruerse in his lippes, and is a foole. 1639 S. DU VERGER tr. Camus' Admim. Events 12 Who for a kingdome would not have blemished her integrity. 1795 Gentl. Mag. 543/1 In integrity of heart and uprightness of intention he was excelled by few. 1850 M'COSH Div. Govt. III. i. (1874) 276 Mankind do, in fact, trust in a person known to be of thorough integrity, that he will always be upright.”

144
unified concept (since the literal meaning of integrity is less complex than the more derivative concept) and we would have some hope of making sense of the common yet under-studied notion of a unified, coherent life as a desirable goal for a moral agent. Second, I believe the more primitive sense of integrity provides the bones of a more interesting and revealing cluster of insights into human agency than the more refined “moral sense” of the term. Finally, I do not mean by my choice of the broader sense of integrity to exclude the narrower “moral sense”; on the contrary, I believe the broader sense provides the scaffolding for a theory of integrity that both affirms the narrower sense, and sheds light on its relation to human agency more generally. My account aims to show that the formal (literal) and substantive (moral) senses of integrity are integrally related rather than free-standing from each other. Whether these claims are ultimately sustainable cannot be judged until I have set out my theory of integrity in more detail.

I take the notion of *wholeness* to be at the conceptual core of integrity. Implicit in this concept are three elements: that which has integrity is

(a) *undivided.* That is to say, the object of integrity has some principle of internal unity: it is not at odds with itself; there is something about it that prevents it from breaking apart.

(b) *complete.* The object of integrity is not missing any essential parts; it is not mutilated.

(c) *internally well-ordered.* What I mean by this is that a series of objects do not form a “whole” or “integer” unless there is some principle by which they are *all* connected, somehow, to every other member of the set. There must be a single “order,” so to speak, which can make sense of calling a multiplicity a “whole.” This could be as
simple as a shared property or structure (e.g. the members of a swing club all want to learn swing), or as complex as a shared goal which all the parts contribute toward in different ways (think of the integrity of an organism such as the human body regulated by the goal of physical health). A whole would be impaired, and in a sense, would no longer be a “whole,” if one of its members no longer contributed to the shared goal (cancerous cells in a body, for example), or one or more of the members ceased to possess the shared property (e.g. some of the members of a swing club no longer want to learn swing).  

Now, applying this basic notion of integrity to a person, we are referring to the unity, oneness, or wholeness, of a person, and all that this entails: a person is (a) not divided within himself; (b) is complete—does not lack any essential feature of personhood. Finally, (c) a person of integrity has some internal principle of order that brings his different parts into some intelligible relation with each other. A lot more needs to be said about this three-fold characterization of personal integrity. However, since this is only meant to be a preliminary exploration, I will defer further explanation for my full account, which I offer in section §4.5.

§4.2 The Subject of Integrity

Now, before I enter into a more detailed analysis of the concept of integrity, I need to say something about the concept of a moral agent, since it is moral agents to

---

230 My account draws on an illuminating discussion of “wholes and unities” in chap. 8 of Jeffrey Blustein, Care and Commitment: Taking the Personal Point of View (New York and Oxford: Oxford University Press, 1991). But I disagree with Blustein that personal integrity is simply a matter of “unifiedness of a person’s commitments,” since the parts of a person’s life that require integration extend beyond the person’s commitments to their desires, actions, and beliefs, as we shall see later on.

231 From now on I will use the term “personal integrity” as equivalent to “integrity of the person.” As such, I do not mean to distinguish personal integrity from, say, political or moral integrity.
whom we ascribe personal integrity. I do not believe any ambitious theory of personhood or agency is required to underpin the notion of integrity as I understand it. However, we do need to know, roughly, what sort of entity we are ascribing “integrity” to. And if integrity involves a relation of parts to whole, what are the relevant “parts” of the possessor of integrity? To start with the first question, my concern here is to identify certain distinguishing features of entities to whom we can sensibly ascribe “integrity.” Whether these features add up to an “essence” or fully constitute the “nature” of a moral agent, is not my concern here. My rough account of moral agency is less a matter of defining the essence of moral agents, than a matter of picking out some important properties that a subject must possess in order to make him/her a candidate for the ascription of “integrity.”

Firstly and most obviously perhaps, the ascription of personal integrity is not a matter of pointing out some natural property of the subject that is wholly independent of its actions; on the contrary, personal integrity is at least partly the outcome of a subject’s actions. And not just any actions can be the source of personal integrity; they must at least include actions animated by intelligence and a rational will, what traditional medieval philosophers called “human acts” (acti humani)\(^{232}\). By “rational will,” I mean a desire that is partly constituted by a rational assessment of its object, rather than simply determined by its object. In other words, we ascribe integrity to creatures who can choose how to act, and whose actions can be controlled in advance by their judgments about the value of the various possible actions available to them. Similarly, we assume that candidates for integrity can recognize deficiencies and merit in their own actions both

\(^{232}\) See for example Aquinas’s *Summa Theologiae*, I-II, a. 6, Q 1, “Whether there is anything voluntary in human acts?”
during and after the fact; in other words, their own actions are not just events which they experience helplessly (such as an earthquake), but reflections of the values they have chosen to act on. In other words, potential candidates for integrity can stand back from their own actions and assess them as worthy or unworthy, sensible or foolish, praiseworthy or blameworthy: that is to say, they can make reflexive judgments as to the worth of their own actions, whether retrospective, actual, or future-oriented.

Finally, we not only hold candidates for integrity responsible for the pattern of actions that flow from their desires; we also hold them responsible for the pattern of desires itself, at least insofar as those desires are cultivated or discouraged over time by the agent’s choices. Thus, a person who acts on the desire to steal is responsible not only for the act of stealing, but for contributing, through his actions, to the psychological dominance of the desire to steal, just as a person who embezzles a company’s profits is responsible not only for the act of embezzlement but for developing, through his choices, an uncontrollable and morally paralyzing desire for material wealth. In short, we expect candidates for integrity to order and cultivate their desires, to the extent that it is within their power, according to some evaluative criteria. In other words, desires are not automatically morally innocent or neutral—they may be healthy or unhealthy—and agents are responsible for feeding or discouraging various desires as well as voluntarily acting (or not acting) on them. “I did it because I wanted to,” with no further qualification, is not a justification for an action.

So far, I have argued that any subject to whom we can sensibly ascribe personal integrity must be a moral agent, and I have enumerated some of the necessary conditions for moral agency. To be a moral agent, on this view, is to have the capacity to (a) use
intelligence and rational will to assess the value or worth of an act and choose to act in accordance with (or against) that assessment; and (b) to use intelligence and rational will to assess one’s desires not only in terms of their relative urgency (desire A is to be satisfied before desire B before desire C), but in terms of their relative and absolute worth or nobility (desire A is more noble or worthy than desire B, desire C is ignoble and to be rejected, etc.).

Finally, (c) as a consequence of a and b, moral agents are, special circumstances aside, generally held responsible for their actions and for their desires insofar as they come under their rational control, meaning that they are held blameworthy for their bad or harmful actions, and praiseworthy for their good or beneficial actions. In short, the subject of integrity is a responsible moral agent. Since we do not ascribe agency in this sense to babies, or to humans who lack the requisite faculties (e.g. somebody who is severely brain-damaged), the class of humans to whom we can ascribe personal integrity is substantially smaller than the class of all humans.

The second question I raised about the subject of integrity is, what are the “parts” of this subject which we expect to be more or less “integrated”? The subject, as we have seen, is a (responsible) moral agent, by which we mean a person whose actions are guided by intelligence and rational will, and who can assess and act upon the worth of

---

233 The ability to form desires about our desires, and not be completely at the mercy of our immediate desires, is one of the most important distinguishing features of persons, according to Harry Frankfurt’s “The Concept of the Person.” Building on this concept of personhood, Charles Taylor, in “What is Human Agency?” argues that some second-order desires involve “strong” evaluation where we are not merely concerned with the satisfaction of certain desires (this is “weak” evaluation), but with the intrinsic worth or nobility of our desires. For example, it is one thing to defer a cream bun in order to enjoy a swim (here, we are concerned with the efficient satisfaction of existing desires and their intrinsic worth is not under scrutiny), and quite another to choose to marry someone rather than joining the army (here, it is precisely the comparative worth or nobility of different desires that is at issue).

234 Naturally, certain qualifications apply. I am certainly not claiming that an agent is blameworthy for all of his harmful actions, but he is certainly blameworthy for many of his harmful actions over the course of a life.
possible actions and the intrinsic worth as well as the strength of his desires. This characterization of the moral agent provides us with a preliminary answer to the question as to the parts of a person to be integrated: the parts are the parts of the person that partially come under the control of the intelligence and will, namely, a person’s desires, evaluations, beliefs, actions, and dispositions.\textsuperscript{235} It is important that the parts to be integrated at least partially come under the control of intelligence and will, otherwise a moral agent could not be held responsible, at least to some extent, for his own integrity (or lack thereof).

§4.3 Existing Accounts of Integrity

Since Bernard Williams’s now famous argument that utilitarianism undermines the integrity of the moral agent\textsuperscript{236}, there has been a proliferation of articles and books treating integrity as an attribute of moral agents, sometimes on its own terms, other times as implicated by some other issue (for example, in terms of the relation between integrity and utilitarianism, or as part of a broader account of human agency).\textsuperscript{237} I do not propose

---

\textsuperscript{235} Dispositions, though they include habitual desires, are not exhausted by them. For example, if I am disposed to feel angry when an injustice is committed, this presupposes a habitual desire to see justice done, but is not reducible to a habitual desire.


to provide a rich and detailed taxonomy of the various conceptions of integrity advanced so far.\textsuperscript{238} Instead, for my purposes, a more modest taxonomy is required. The main purpose of this limited taxonomy is not to capture all the nuances of the discussion but to highlight the central features of extant accounts of integrity that I find implausible, and the central features that I see as valuable or insightful. It turns out that this winnowing process requires a relatively simple, albeit somewhat artificial, classification of existing accounts. I will provide some preliminary justifications for my selective adoption of aspects of existing accounts as I critique each category of integrity (§4.4), but the fuller justification will be the positive account of integrity that emerges from the critique (§4.5).

Most accounts agree that some degree of integrity in a person’s life is both necessary and desirable. There is broad agreement that a moral agent who possesses integrity not only has a certain order among his desires, evaluations, beliefs, and actions\textsuperscript{239} at a given moment, but has a stable disposition to act in ways that preserve or

\begin{footnotesize}
\textsuperscript{238} One of the more influential taxonomies along these lines can be found in Calhoun, "Standing for Something," and is elaborated in Cox, Caze, and Levine, \textit{Integrity and the Fragile Self}, chap. 1, “Views of Integrity.”

\textsuperscript{239} Partly on account of the modern bias against treating the emotions and settled dispositions as under the sway of reason (for a good argument to this effect, see Hursthouse, \textit{On Virtue Ethics}, chap. 5), modern
\end{footnotesize}
promote his integrity (whether that outcome is something he consciously intends or not). However, philosophers differ dramatically in the particulars of their accounts of integrity and their assessment of its overall importance in the moral life, how if at all it should be deliberately pursued, and whether it is unqualifiedly a *virtue* rather than just a necessary but insufficient condition for virtue to exist. For example, if we view integrity as an intrinsic good, then the stable disposition to act with integrity might be considered unqualifiedly a virtue; but if integrity is only conditionally or contingently good, then the stable disposition to act with integrity is only conditionally a virtue. There are different ways of categorizing accounts of integrity, but I believe the most fundamental question that marks out the orientation of different accounts is, “what is the relation between *integrity*, and moral goodness *simpliciter*?” Accordingly, I group accounts of integrity into five categories or families:\(^{240}\):

accounts of integrity do not tend to treat dispositions as telling for or against a person’s integrity. Thus, in order to capture the nature of existing accounts I do not include dispositions in my taxonomy, although as we will see, I do believe they are an important element of integrity.

\(^{240}\) My taxonomy is very different from the one suggested by Cheshire Calhoun, in “Standing for Something” (1995) and appropriated by Cox, La Caze and Levine in *Integrity and the Fragile Self* (2003). Calhoun groups accounts of integrity into three families: the “integrated self,” “identity,” and “clean hands” pictures of integrity. On the “integrated self” picture, integrity consists in the integration of one’s various desires, evaluations, and actions, into a coherent whole, such that they do not contradict one another. The moral agent is thus not “divided” within itself. On the “identity” view, integrity consists in fidelity to one’s most central or “core” commitments, those commitments that define who one is—what some call “identity-conferring commitments.” Finally, on the “clean hands” view, the agent of integrity has certain normative commitments that he or she will never violate, no matter the cost. The person of integrity keeps his hands “clean.” The reason I do not adopt Calhoun’s classification scheme is because I do not believe it adequately illuminates what is really at stake in debates about integrity. To start with, I see the “integrated self” and “identity” pictures of integrity as having such a close mutual affinity that they are, to a large extent, the same sort of account looked at from different angles. At bottom, both views emphasize the internal coherence of an agent’s commitments, including the agent’s scheme of priorities, and therefore both views may be considered within my taxonomy as stabilized formal accounts of integrity. Finally, whereas the third element of Calhoun’s scheme of classification, the “clean hands” picture of integrity, emphasizes the unconditionality of moral commitments in a person of integrity, my partially substantive category of integrity is wider in scope, taking in a range of substantive limitations on an agent’s commitments, whether those limitations are prudential or absolute. Thus, if we believe that truthfulness is a basic feature of a person of integrity, this is a partially substantive account, even if we do not believe that integrity requires us to tell the truth *in every conceivable eventuality.*
(i) *Purely formal accounts* essentially demand internal consistency within the *form* or structure of an agent’s desires, actions, beliefs, and evaluations. On the purely formal account, integrity itself is compatible both with good and evil desires, actions, and evaluations. Therefore, while it may enhance a substantively good life, it may also exacerbate a substantively evil life. Consequently, in and of itself, integrity is neither a virtue nor a vice: it is just a formal property of an agent’s commitments. On the purely formal account, somebody may be committed to evil causes or principles, and they may adopt principles of expediency or even exempt themselves from moral rules when the rules stand in the way of their desires, once such exemptions are sanctioned by the agent’s own principles or beliefs. Few philosophers are willing to embrace an account of integrity so bereft of normative constraints as this one, which is compatible with capricious and erratic behavior provided such behavior is sanctioned by the agent’s commitments.

(ii) *Stabilized formal accounts*, like purely formal accounts, demand internal consistency among an agent’s desires, actions, beliefs, and evaluations. However, unlike purely formal accounts of integrity, this account rules out unstable or highly changeable evaluations, as well as insisting that evaluations apply a consistent standard of action, relatively independent of changes in the agent’s desires or circumstances. This notion of stable, consistent evaluation is captured by the term “principled commitments.” On the stabilized formal account, an agent of integrity must not only possess internal consistency among his desires, actions, beliefs, and evaluations; evaluations must embody consistent, stable, and enduring standards of action that are not at the mercy of the agent’s shifting desires and circumstances. While stabilized formal accounts do filter out extreme cases of
wantonness, capriciousness, and superficiality, they are also consistent with the adoption of a “principled commitment” to evil in general or to some particular evil cause.\textsuperscript{241}

(iii) \textit{Subjectively moral formal accounts}, like stabilized formal accounts, demand internal consistency among an agent’s desires, actions, beliefs, and evaluations, as well as stable and enduring evaluations. However, subjectively moral accounts also demand that the agent \textit{intend} to do what is right or good, and believe that his actions satisfy that intention. In other words, the overarching goal of the agent must be to do what is \textit{good} or \textit{right} rather than what is evil or wrong, and he must believe that his actions are conducive to that goal.

(iv) \textit{Partially substantive accounts} have all the features of subjectively moral formal accounts, namely, (a) they demand internal consistency among an agent’s desires, actions, beliefs, and evaluations, (b) as well as stable and enduring evaluations, and (c) the intention to do what is right or good. However, in addition, unlike the three formal accounts considered thus far, partially substantive accounts of integrity place some \textit{substantive constraints} on the sorts of actions that are compatible with personal integrity. Integrity-preserving desires, actions, beliefs, or evaluations must not violate certain objective moral standards. This does not mean that integrity-preserving desires, actions beliefs, or evaluations are necessarily \textit{good} or \textit{right} (although of course this is always a

\textsuperscript{241} Williams’s notion of fidelity to one’s “ground-level projects” is what I am calling a stabilized formal account of integrity. See Williams, “A Critique of Utilitarianism.” Frankfurt’s understanding in “Freedom of the Will and the Concept of the Person” of the integration of first- and second-order desires as central to human agency also provides the basis for a stabilized formal account of integrity. Another example of this approach to integrity is Rawls’s, in TJ, 455-456: “we should note that at times of social doubt and loss of faith in long established values, there is a tendency to fall back on the virtues of integrity: truthfulness and sincerity, lucidity and commitment, or, as some say, authenticity…Now of course the virtues of integrity are virtues, and among the excellences of free persons. Yet while necessary, they are not sufficient; for their definition allows for most any content: a tyrant might display these attributes to a high degree…”
possibility), but it does mean that they cannot be egregiously evil or wrong. In other words, partially substantive accounts do not guarantee the goodness or rightness of actions they deem compatible with integrity, but they do place constraints on how far an action can wander from goodness towards evil. This sort of constraint is sometimes expressed by the thought that actions should be minimally “decent” or “reasonable,” and the source of standards of decency and reasonableness may be located in the minimal conditions for living at peace with other persons (e.g. respect private property, do not gratuitously injure your neighbor, and so on).\textsuperscript{242}

(v) Fully substantive accounts are at the opposite extreme to purely formal accounts on the spectrum of formality. Fully substantive accounts share the demands of accounts (i) through (iv), but they define integrity in such a way that a life of integrity is existentially coextensive with, though not conceptually identical to, a good life. That is to say, insofar as a person attains integrity, he approaches what is (objectively) good or right, and insofar as he loses integrity, he distances himself from what is (objectively) good or right. On this account, the person who has fully attained integrity has fully attained the human good. Even if no such person actually exists, the claim that integrity demands all-round moral goodness is an important one, that is altogether at odds with formal accounts which suggest that any coincidence of integrity with moral goodness is at best a matter of probability (he who commits himself to what he believes are morally good goals seems at least more likely to do good than he who commits himself to what he

\textsuperscript{242} This is a particularly popular view of integrity in the literature. Representatives of this sort of view on the less demanding end of the spectrum include McFall, "Integrity," ; Blustein, Care and Commitment ; Cox, Caze, and Levine, Integrity and the Fragile Self ; Taylor, Pride, Shame, and Guilt . On the more demanding end, we find Graham, "Does Integrity Require Moral Goodness?,”
believes to be morally evil goals); or perhaps even a purely contingent matter (on the purely formal account, for example).243

§4.4 A Critique of Existing Accounts

So far, I have explicated five views of integrity, from the fully formal to the fully substantive. Now, I would like to highlight several limitations affecting discussions of integrity across all five views. The first concerns more formal accounts of integrity, which purchase conceptual unity and simplicity at the cost of relevance to and insight into everyday experience. These views reduce integrity to a sort of formal consistency among desires, evaluations, beliefs, and actions, but argue that integrity is compatible with almost any action, good or evil. Thus, more formal accounts, in their failure to account for the normative dimension of the concept of integrity, are radically incomplete.

Many accounts, however, are partially substantive, and though they do succeed in establishing some connection between integrity and objective goodness, they nonetheless share a common defect, namely, that they lack conceptual unity or integrity (in the more limited, formal sense). What I mean by this is that they offer an aggregate of properties, qualities, or conditions, that are thought to justify the ascription of integrity to a person, but the aggregate itself is just that: an aggregate with little or no conceptual core. This leaves theories of integrity floundering as they attempt to incorporate various intuitions latent in ordinary uses of the concept, without any unifying concept or phenomenon to guide the inquiry. Without a core concept, a theory based on ordinary language analysis

243 The account that comes closes to being fully substantive is Ashford, "Utilitarianism, Integrity, and Partiality."
easily descends into a *survey* or collation of ordinary language uses of a concept, grounded in little more than the need for logical consistency among uses.

What we need, however, is not an analysis of ordinary language usage, but a *theory* of integrity that, while not entirely disconnected from ordinary language, is capable of discerning a moral core to the concept of integrity—a core that can both illuminate our moral experience and critically distance us from conflicting and misleading uses of the term as it occurs in ordinary discourse. What we need is not a survey of ordinary uses of the term “integrity” (this is the sort of job a dictionary can do), but a philosophical defense of one conception\textsuperscript{244} of the term and its superiority over rival conceptions, which cannot be established based purely on ordinary language usage of the term, since this is diverse and conflicted, but based on an interpretation of our moral experience. The point is not to ignore or disregard ordinary language, but to realize that it constrains rather than uniquely determines the shape of the concept of integrity. To put it another way, ordinary language use is compatible with a range of concepts of integrity, but we must endeavor to offer a unified interpretation of our moral experiences in order to justify the adoption of one concept rather than another. Thus, if we developed a concept that seemed utterly disconnected from most uses of the word, we would be coining a new concept rather than analyzing an existing one. But a conception of a term (e.g. my conception of integrity) may bear a family resemblance to many of its uses, discriminating in favor of some and against others, while still remaining at bottom the same *concept*. The point of ordinary language analysis, then, is not to discover the “true” meaning of a word, but to better bring into relief some of the common understandings of

\textsuperscript{244}Dworkin builds his theory of legal interpretation around the distinction between the *concept* of law and its rival *conceptions*. See Ronald Dworkin, *Law's Empire* (Cambridge, Mass.: Belknap Press, 1986).
the term, so that these can then be compared with experience, and subjected to a more probing philosophical critique. In short, ordinary language analysis is the handmaid of philosophical argumentation rather than the other way round.

Closely related to the problem of conceptual fragmentation, most theories, including the more formal ones, fail to show any interesting relation between the form of agency required by integrity and the substantive demands it places on agency. The relation is nearly always simply additive, e.g. integrity requires both coherence among values and the “right” sorts of values. But while this may capture the way the word is frequently used, it is not a particularly interesting or compelling theory, since it fails to show how coherence and the correctness of values are both reflections of the same concept. It just says that they are, but it fails to answer the question, what is the connection between form and substance? How is it that they constitute a single concept or virtue (namely, integrity)? I believe that a more conceptually unified or integrated theory has the potential to significantly mitigate the ad hoc quality of existing accounts of integrity, as well as offering more interesting and lucid insights into why integrity is such an important virtue in the moral life.

Of the five accounts I have listed, it should be evident by now that I reject both the purely formal and stabilized formal accounts as neither successfully resonating with ordinary usage of the concept of integrity, nor shedding light in a sufficiently interesting way on moral agency as we experience it. So the three most plausible contenders are (a) the subjectively moral formal account; (b) the partially substantive account; and (c) the fully substantive account. All three of these accounts establish a connection, of varying
degrees of robustness, between the possession of integrity and the pursuit of good and right ends.

The fully substantive account, which judges the person of full integrity to be ontologically or practically identical with (though conceptually distinct from) the fully virtuous person, does not seem plausible, for at least two reasons: first, it has broken so decisively with ordinary usage that it is hard to see how it could be considered an analysis of ordinary moral experience. The word is hardly used in the fully substantive sense by anyone, except perhaps by some moral philosophers. For example, most of us have no hesitation in ascribing a high degree of integrity to people whose moral views diverge substantially from our own, and—assuming we believe our own views to be true—this means that integrity is overwhelmingly perceived as compatible with objectively mistaken or bad beliefs and actions. This is not to say that the fully substantive interpretation of integrity is manifestly wrong or indefensible: but it is to say that what started out as an enterprise in giving a plausible account of a familiar feature of our moral experience, integrity as a more or less praiseworthy quality of persons, has become instead an effort to put forward a theory of integrity that has very little connection to the value of integrity as understood and lived by ordinary people.

What began as an effort to refine our understanding of a moral term of common currency ended up as an effort to defend an ambitious interpretation of the moral life according to which full personal integration corresponds neatly to full or unqualified personal goodness or virtue. I do not reject this interpretation of the moral life, but I do believe it requires a much more elaborate defense than I am prepared to offer here, and it is a project distinct from the interpretation of integrity as we commonly understand the
term. In any case, a more modest and pedestrian notion of integrity does not entail the
outright rejection of the more ambitious interpretation, and is sufficient to capture the
phenomena I have in mind. Admittedly, the more ambitious interpretation does not seem
plausible, and would require a very elaborate defense to become plausible, but it may turn
out, in light of such a defense, that there are two distinct conceptions of integrity, which
simply describe different aspects of moral agency, and complement rather than contradict
each other.

So much for the fully substantive account of integrity. Now, what of the
subjectively moral formal account? Insofar as this account insists that the agent must
intend to do what is good and take the steps deemed (by the agent) necessary for that end,
this seems to capture quite well what most people mean by integrity. However, in failing
to draw a closer connection between the intention to do what is right or good, and the
substantive goodness of the person and his acts, the account is unsatisfying and fails to
explain why we balk at the thought of attributing integrity to a well-intentioned Nazi or
Stalinist. In short, the subjectively moral formal account fails to explain the relation
between the formal aspect of integrity, which suggests coherence and good intentions,
and the substantive aspect of integrity, which suggests personal virtue and good action.
Any account that disregards or fails to address either component of integrity is
inadequate.

The partially substantive account has the virtue of insisting on at least some
substantive constraints on the actions of the person of integrity. However, this account
seems a bit ad hoc because it is not clear what the connection is between the concept of
integrity as a formal character trait and the “reasonableness” of the commitments of the
person of integrity. Doubtless, this identification may coincide to some extent with ordinary usage of the term, but it does not explain or justify it. What we require is not the assertion of a link between form and substance, but an explanation of the link. As will become clear shortly, I espouse a version of the partially substantive account that aims to connect—or, if you will, integrate—the form and substance of the virtue of integrity more cogently than other accounts.

§4.5 Integrity as a Family of Virtues

§4.5.1 The Relationship between Integrity and Virtue

The main challenge facing any interpretation of integrity is to be both sensitive to ordinary usage of the term and capable of offering a theoretically cogent and discriminating account of this quality we frequently attribute to moral agents (including, perhaps, ourselves). What I propose to do here is to begin to sketch an account of integrity that meets this challenge. This will require addressing some issues—most notably, whether or in what sense integrity is a virtue—that could not be adequately addressed by the general taxonomy I presented earlier. In addition, in light of the political philosophic concerns of my project, it will be necessary to flesh out my account some more to show that citizens may, pace Rawls, reasonably pursue integrity—at least as I understand the term—in a wholehearted or unreserved way. I do not see this as a matter of setting out my account first, and then defending it against liberal objections; rather, I see it as a matter of setting out a preliminary and somewhat crude account of integrity and then refining it and developing it further in response to the concerns motivating some standard liberal objections. It is crucial to grasp this point, otherwise it may appear that I
am offering a formal account of integrity first and only subsequently sneaking in more substantive claims. Whereas in fact, the substantive claims are integral to the account, and the account stands or falls with them.

As I suggested in the previous section, I favor an account of integrity that is partially substantive, in the sense that it places genuine, objective moral demands upon a person’s actions, yet does not neatly correspond to, or entail, virtue *simpliciter*. In this respect my account is similar to other partially substantive accounts. However, I hope to forge a much more compelling link between the formal aspect of integrity, namely the good intentions and coherence of a person’s life, and the substantive aspect, namely the objectively good character-traits and acts instantiated in a person’s life.

In order to understand the connection between the formal and substantive aspects of integrity, we must inquire in what sense, if any, it is a virtue. But to tackle this question, we must begin by asking, *what is human virtue?* A good place to start is Aristotle’s famous definition of virtue: having established that virtue is neither a capacity nor an emotion but a disposition, Aristotle goes on to specify what *kind* of disposition it is:

> It must then be premised that all excellence has a twofold effect on the thing to which it belongs: it not only renders the thing itself good, but it also causes it to perform its function well. For example, the effect of excellence in the eye is that the eye is good and functions well; since having good eyes means having good sight … If therefore this is true of all things, excellence or virtue in a man will be the disposition which renders him a good man and also which will cause him to perform his function well.245

---

Following Aristotle, I want to suggest that human virtue—the excellence distinctive of a human person—is the excellence of those functions that set human beings apart from other animals, which are, as we have seen, the capacity to know and will, both in relation to other things, other persons, and oneself. To be a *good person*, then, must be to know and will *well* in relation to other things, other persons, and oneself. To *know well* is to habitually and by disposition know what is *true* and what is *good*, whether in the theoretical sphere or in the practical sphere. To *will well*, is to habitually and by disposition desire and *do* what is good. Having a *good will* pertains to the practical rather than the theoretical sphere, while having *good knowledge* has both theoretical and practical dimensions.

So far, this seems rather abstract, perhaps. But it is not as abstract as it sounds. We use the word “good” all the time, and know very well what we mean, once we are given a context in which to interpret it. I am not suggesting that we always (or even nearly always) agree on what is good; but I am suggesting that human goodness and excellence is much less problematic and controversial than it is now fashionable to think. Human goodness is not constituted merely by an isolated good deed—thus, the serial killer who out of compassion gives his victim a morsel of food is not thereby a “good person”; but neither is human goodness constituted by a single type of abiding disposition to do good deeds—thus, the emotionally unresponsive father is not considered a good person just because he is habitually conscientious and honest with his family at home and his clients at work. Rather, human goodness is constituted by a complex ensemble of

---

246 In fact, there is no exact equivalent of the modern concept of will in Aristotle, though he does have the concept of choice, and personal responsibility for actions that come under the sway of reason (see especially NE 5, viii, 1135a16 ff). It was not until later Christian thinkers, such as Augustine and Aquinas, that the concepts of intellect and will, and indeed the modern concept of person, emerged more fully.
abiding dispositions to act well in all or most of life’s situations. To be a good human person, in the fullest sense of the term, is to possess the full complement of human virtues, or at least those virtues or dispositions of character that are required in order to respond well to the events, persons, and objects one encounters, including oneself.

We have already identified the functions distinctive of human beings, those functions that raise them above non-human animals, as the capacity to know and to will in a reflective and responsible way. Just as a good non-human animal is identified according to the degree to which it has realized its distinctive functions well, whether mating, eating, or physiological development, a good human being or person is identified according to the degree to which he or she has realized his or her distinctive functions well, viz. knowing and willing.\(^\text{247}\) Although we have no comprehensive and infallible body of principles telling us what good knowing and good willing are, we can readily identify the difference between a stunted, maldeveloped human, and a flourishing, healthy human, not just in physical terms but in psychological and moral terms. We know that willing the destruction of one’s children is callous and may even be indicative of a deep psychological disorder; just as we know that extending a helping hand to one’s friends in need is a sign (though by no means a decisive or sufficient sign) of good character and a healthy psyche. Though we may disagree fundamentally on what the right thing to do is in such-and-such a situation, most of us are agreed that dispositions to be honest, to welcome strangers, to extend assistance to friends in need, to be sincere with ourselves, to tell the truth, to finish a task we have begun, and so on, are virtues and not

vices, though in particular cases even these dispositions need to be corrected by other
more pressing or pertinent dispositions (e.g. diligence in completing a task becomes a
vice if not displaced by compassion for a sick child). Each of these virtues clearly
constitutes a perfection of the human will, since they all assist a person to develop
healthy or good desires vis-à-vis the persons and objects around him or her. They are also
perfective of the human intellect, since they make knowledge of the correct thing to do
somehow “second nature,” something that often (though by no means always) requires
little reflection.

Similarly, though we may disagree in many particular instances on whether a
theoretical proposition or claim is true or false, we can agree that the habitual capacity to
understand difficult concepts, to give plausible explanations of natural phenomena, to
complete a thorough scientific investigation, to quickly grasp and articulate the essence
of an author’s thesis, to make interesting connections between different bodies of
knowledge, and so on, are dispositions that perfect the human intellect, i.e. intellectual
virtues or excellences of the human person.

Now, we may distinguish with Aristotle between imperfect virtue, a disposition
which makes a person good after a manner or in a partial, less than perfect, fashion; and
perfect virtue, or virtue without qualification, a disposition which makes a person good
simply. No single virtue or subset of virtues, in isolation from the others, can render a
person good without qualification, since only the person whose every disposition
conduces to his excellence can be considered good without qualification. Hence, on
Aristotle’s view, while we may possess some virtues imperfectly while lacking others, no
single virtue can be possessed perfectly without possessing all other virtues perfectly.
This is the doctrine that has become known as the “unity of the virtues.” If integrity is a specific virtue or set of virtues, and if we accept the account of virtue I borrow from Aristotle, then integrity, like any other virtue, may be possessed either perfectly or imperfectly. And in practice, given the extreme difficulty (impossibility?) of finding people who instantiate all the virtues perfectly, in almost all cases (barring truly extraordinary cases such as the perfect saint), integrity will be an imperfect virtue, i.e., a virtue that renders a person good after a manner or according to some partial measure of goodness. Also, like any virtue, integrity will be a matter of degrees rather than an on-off quality: just as one person may be more or less virtuous than another, one person may have more or less integrity than another.

But the question remains, can we describe integrity as a virtue, and still remain fairly close to ordinary usage as well as yielding interesting insights into our ordinary moral experience? I believe we can, but it is not a straightforward matter. Unlike virtues such as courage, magnanimity, justice, and patience, integrity is notoriously difficult to identify with one specific sort of excellence. It is inexorably identified with a range of virtues, all related in some way to the agent’s pursuit of goodness, such as truthfulness, honesty, sincerity, steadfastness, responsibility, open-mindedness, humility, and conscientiousness. Granted that these are all excellence-inducing dispositions, the challenge for a theory of integrity is to explain how a single virtue-term, “integrity,” can encompass all of these dispositions without collapsing into an ad hoc laundry-list of virtues.

---

248 For one eloquent summary of the essentials of this doctrine, see Hursthouse, *On Virtue Ethics*, 153-157.
I believe the formal definition of integrity we started out with, namely, the coherence of an agent’s beliefs, actions, evaluations, desires, and dispositions, provides the key to a unified account of integrity, provided we grant that (a) the agent’s highest-order desire is to do the good, and (b) that desire is effective and stable enough to dynamically structure and re-structure the agent’s moral life (beliefs, actions, evaluations, desires, dispositions) over time. By this I mean to exclude as a candidate for integrity the agent who has a highest-order desire to do what is good (“I really do want to want to do what is good, in spite of all of the desires that get in the way”) but whose other (inconsistent) desires consistently overpower it (this would characterize the “wanton”). An agent whose desire to do what is good is rarely if ever effective (i.e. issuing in the appropriate action), is not a person of integrity.

What integrity calls for, in short, is that the desire to do what is good be generally both *highest* among an agent’s desires and *effective and stable enough* to dynamically shape or order an agent’s moral life (beliefs, actions, evaluations, desires, dispositions) over time. Integrity, then, is in practice not merely a formal relation of unity among one’s beliefs, actions, evaluations, desires, and dispositions, but a specific disposition or set of dispositions that reflects a dominant, stable, and effective desire to do what is good or right. By effective, I do not mean that the agent actually accomplishes what is objectively good: rather, I mean that the agent takes appropriate steps to discern what is objectively good\(^{249}\) and orders (or reorders) his moral life (beliefs, actions, evaluations, desires,

\(^{249}\) This epistemic dimension of integrity, what we might call “epistemic responsibility” for short, which is closely related to conscientiousness (though, unlike conscientiousness, it does not directly connote specific responsibilities to others), was brought home to me in particular when I read Carter, *Integrity* (see, for example, 22).
dispositions) in accordance with that determination.\footnote{I am not suggesting that the agent who fails to order his moral life *perfectly or completely* in accordance with his determination of the good is seriously deficient in integrity. But I do believe that the high benchmark for the person of full integrity—a benchmark which no fallible human, to our awareness, meets—is a perfect ordering of one’s beliefs, actions, evaluations, desires, and dispositions, in accordance with one’s responsible discernment of the good.} That is to say, we can distinguish an epistemic as well as a practical dimension of integrity. Usually, the epistemic aspect involves practical steps such as seeking out information and advice from reliable sources, as well as making a practical judgment about what the best course of action is, here and now. Clearly, the process of learning about the good is an inherently practical affair. Therefore, it might be more helpful to distinguish between the immediately knowledge-oriented practical dimension of integrity and the immediately action-oriented practical dimension.

Before I discuss the virtues of integrity, I need to say a bit more about my claim that the person of integrity seeks to do what is objectively good. It is important to say this in order to guard against one misreading of integrity, according to which the person of integrity simply seeks to follow his or her opinion about the good. If all I were accountable to were my opinion, then I could rest content with whatever opinion I happened to hold, and would not feel any impulse to revise my opinions in light of new evidence and facts. We could say that I seek to follow my opinion, *providing it is well-founded*, but this already carries me beyond my opinion to a world of moral experience and values that pre-exists or exists independently from the process of discernment and practical judgment. We could not make sense of the commonplace moral experience of struggling to discern what is right or good without positing an objective realm of values.

\footnote{I say, *immediately* knowledge-oriented because all practical knowledge, i.e. all knowledge of the good in the practical sphere, is ultimately oriented towards action rather than knowledge.}
or a moral reality that subjective opinion responds to. Some, such as Richard Rorty, might want to dismiss this experience as inherently delusional. I want to take it on its own terms, and analyze it more in a phenomenological than a skeptical spirit.

§4.5.2 The Virtues of Integrity

So far, there has been a fairly abstract and overly formal feel to my discussion of integrity. The reader may wonder what integrity amounts to for real, live, breathing human beings who do not routinely divide themselves up into beliefs, actions, evaluations, desires, and dispositions when they act. I hope the relevance of this formal account of integrity will become clearer as we consider the specific family of dispositions we naturally associate with integrity. Recall that integrity entails that (a) an agent’s desire to do (what is objectively) good is higher than all other desires, and (b) this highest desire is stable and effective in the sense that the agent is in the habit of taking appropriate steps to determine what is objectively good and acting upon that determination. We would do well to note immediately that “taking appropriate steps” to determine what is objectively good, though it certainly enhances the credentials of the determination, is not the same as correctly determining what is objectively good. Therefore, there is no neat

252 In other words, our moral attitudes and behavior seem to presuppose some version of moral realism, which I understand, roughly, as the view that moral principles, norms, and values are not validated merely by our affirmation of them, but by virtue of their responsiveness to a reality that exists whether or not we acknowledge or grasp it. Whether we describe this reality as a realm of “moral facts” or as a purpose-driven universe, or in some other way, is not something I need to address here.

253 See, for example, Richard Rorty, Contingency, Irony, and Solidarity (Cambridge and New York: Cambridge University Press, 1989).

254 I do not want to say, “acts always” upon that determination, because it is noticeable that even people we consider highly virtuous suffer from weakness of will and other vices that interfere with acting well. Perhaps we could say that how much integrity a person possesses is partly a function of how often or reliably she acts upon her conscientious determination of the good.
correlation between “acting upon one’s responsible determination of what is objectively
good” and “acting in accordance with what is objectively good, period.” Nonetheless,
they are positively related in interesting ways, as we will see shortly.

**Temperance**

I want to argue that integrity is not a single virtue but a cluster or family of virtues
closely associated with the persistent effort to pursue a good life. What virtues or good-
making dispositions must an agent possess in order to (a) have a persistent highest-order
desire to do what is (objectively) good; (b) take appropriate steps to determine what is
(objectively) good; and (c) order or re-order his moral life (beliefs, evaluations, desires,
dispositions, actions) over time in accordance with that determination? One obvious
virtue required if the desire to do good is to be *effective* and not displaced by other
inconsistent desires is temperance, or the ability to enjoy bodily pleasures with
moderation or due measure. If the desire for bodily pleasure (for example, food, sex,
alcohol, or drugs) becomes too dominant or excessive, either a person’s ability to discern
what is good all things considered, or his ability to *act* upon that determination, is
severely damaged, perhaps even disabled. Thus we may say that the adulterer lacks
integrity when he sets aside his duties to his family and spouse because his desire for
pleasure predominates over his desire to do what is genuinely good (assuming for the
sake of argument that this is his highest-order desire).
Prudence

Similarly, on the epistemic-practical side of integrity, some measure of prudence is surely required, for otherwise how could I know what the “appropriate” steps are in order to determine what is good in such-and-such a situation? The path to knowledge, and in particular the path to moral knowledge, is not mapped out for us in advance. We need to make particular, unrepeatable judgments such as, “From whom should I seek advice on this matter or this dilemma?” and “how much time should I deliberate before acting given the urgency of this matter?” and “what sort of evidence is relevant to the determination of good in this case?” I do not understand prudence in a narrowly self-interested, or amoral, sense. Instead, I follow Aristotle in understanding prudence as the disposition to “deliberate well about what is good and advantageous for [oneself],” understanding “good and advantageous for [oneself]” to mean valuable for the agent in the broadest sense rather than in some narrow, e.g. hedonistic or utilitarian, sense.255

Conscientiousness

Closely related to prudence is the virtue of conscientiousness. I understand this virtue as the disposition to give due consideration to one’s responsibilities, whether simply as a human or as the occupant of some role, and to discharge those responsibilities in as effective and timely a fashion as one can. The conscientious seeker of the good realizes that he is accountable or responsible to others for the outcome of his quest for the

255 Aristotle offers us a simple and lucid definition of prudence in NE 6, v, 1140a125-40: “Now it is thought to be a mark of a man of practical wisdom to be able to deliberate well about what is good and expedient for himself, not in some particular respect, e.g. about what sorts of thing conduce to health or to strength, but about what sorts of thing conduce to the good life in general. This is shown by the fact that we credit men with practical wisdom in some particular respect when they have calculated well with a view to some good end which is one of those that are not the object of any art. Thus in general the man who is capable of deliberating has practical wisdom.”
good (not least to his family and friends, but also his fellow citizens and fellow humans\textsuperscript{256}) and he will therefore bring an attitude of seriousness to the task. This does not imply that the conscientious person never smiles or tells jokes: rather, it implies that he or she puts in an effort to “get it right” that is at least proportionate to the responsibility at issue in the decision (it may be that there is very little responsibility at issue, e.g. I do not deliberate long and hard over whether to eat cornflakes or rice krispies in the morning). A general lack of conscientiousness betrays a very shaky commitment to know and do the good, and thus a lack of integrity.

**Honesty**

That honesty is one of the salient virtues of integrity becomes apparent as soon as we notice that one major obstacle both to the determination of good and to (re)ordering one’s life according to that determination is dishonesty and the closely related vices of untruthfulness and insincerity. Dishonesty is not just a matter of telling untruths, but of failing to confront, or deliberately avoiding the truth about oneself and about reality, carrying on as if oblivious to something that is within one’s field of view but which one would prefer, for whatever reason, to ignore. For example, the heartless landlord who refuses to extend mercy to a delinquent tenant may rationalize her behavior as “fair,” saying to herself, “if I let him off the hook, that would be unfair to the other tenants who have to pay rent.” Indeed, she may suppress or hold at bay a niggling suspicion that her behavior is in fact inhumane or uncharitable. This would be fundamentally dishonest. The reason dishonesty is so damaging to the determination of goodness is that rational inquiry

\textsuperscript{256} For Christians and many other religious believers, conscientiousness is evinced in the believer’s sense of responsibility and stewardship for the gifts he has received from God.
about practical matters, both introspective and dialogic, depends on the ability of all concerned to acknowledge the truth even when it is unpleasant or inconvenient. Seeing what is right or good and acting on that insight depend on an ability to honestly assess one’s own motives, even if that assessment leaves one in an unflattering light.\textsuperscript{257} This is often thought of as “owning up” to one’s actions or taking responsibility for who one is and what one has done. It is not just a matter of acknowledging the good to be done, but acknowledging the evil one has done and being willing to rectify it in whatever way one can.

We might also consider that integrity entails honesty in a more immediate way. For dishonesty in its various forms seems to be constituted by a lack of fit between one’s beliefs, desires, evaluations and actions. Sometimes it is straightforwardly a matter of expressing in one’s speech (an action) values or beliefs that one does not in fact hold. Thus, one’s speech is not reflecting who one is, or better, one’s speech is distorting or contradicting who one is (e.g. what one believes matters most in life). Other times, it may be a matter of forming false beliefs that rationalize rather than explain one’s evaluations and actions. In that case, to the extent that the rationalization is dishonest, one knows “deep down” that the beliefs one is forming are suspect, but does nothing to resolve the contradiction between the two beliefs.

\textsuperscript{257} Of course, honesty has much wider connotations, such as playing by the rules, keeping one’s promises, and so on, but these are, I believe, derivative from the quality of truthfulness, since the reason a thief or a cheat are considered dishonest is because the success of their bad enterprises depends on their ability to deceive others that they do in fact accept the rules of the game or of the society they are members of.
Courage

Courage, no less than honesty, is central to our ordinary understanding of integrity. This may be required both in order to discover the truth (about oneself, about what is to be done) and to act in light of it when doing so is personally costly, dangerous, or even fatal. The paradigmatic cases of integrity involve a persistent quest for the truth, and a noble effort to act upon the truth, in the face of extreme danger. Cases that immediately come to mind are Socrates’s apology before his execution, Thomas More’s refusal to swear allegiance to the Church of England even on pain of death,258 and Martin Luther King, Jr’s proclamation of the equality of all races in a culture in many ways hostile to his claims.259

Perseverance

Perseverance or steadfastness is the disposition to stick with a task and see it through, in spite of obstacles and setbacks such as fatigue, monotony, mistakes, or general unpleasantness, whether it is a relatively minor task such as sweeping the floor, or a lifelong task such as raising a family or becoming a better person. Clearly, in order for somebody to shape and re-shape his moral life (actions, evaluations, beliefs, desires, dispositions) in accordance with (his view of) the good, especially given the array of difficulties and setbacks that can afflict such a task in the case of any normal person, he will require the virtue of perseverance: specifically, a long-range practical commitment to


259 See, for example, Martin Luther King’s famous letter from Birmingham jail, at http://www.stanford.edu/group/King/frequentdocs/birmingham.pdf.
adjust and readjust his life to the good even when the process of adjustment encounters serious setbacks, sometimes on account of “moral luck”\textsuperscript{260} and sometimes on account of his own mistakes or vices. Perseverance in the quest for the good does \emph{not} mean unconditionally holding fast to one’s judgments about the good, since one’s judgments may become problematic in light of new insights or evidence, or they may be found to be unreasonable in the first place. Rather, perseverance means continuously taking steps to ensure one’s judgments are well grounded, and persistently conforming one’s life to one’s judgments about the good just insofar as one believes, in good faith, that they are well grounded.\textsuperscript{261}

**Humility**

Humility does not enter Aristotle’s canon of virtues, nor is it a virtue that is widely discussed by theorists of integrity, but I believe it is central to any plausible understanding of the concept. Humility is the disposition to appraise one’s own character and achievements realistically or accurately even in situations in which one might be tempted to exaggerate or inflate them. It frequently involves a willingness to acknowledge one’s dependency upon others even when that dependency brings into relief

\textsuperscript{260} There has been an explosion of interest in the topic of “moral luck” in recent decades, enjoying the attention of distinguished philosophers foremost among them Martha Nussbaum, \textit{The Fragility of Goodness: Luck and Ethics in Greek Tragedy and Philosophy} (Cambridge: Cambridge University Press, 1986); and Bernard Williams, \textit{Moral Luck} (Cambridge: Cambridge University Press, 1981). Article-length treatments can be found in Daniel Statman, ed., \textit{Moral Luck} (Albany: State University of New York Press, 1993).

\textsuperscript{261} This takes care of the objection that integrity entails a kind of rigidity or refusal to alter one’s practical judgments over time.
one’s own imperfections, weakness, or moral defects. For example, it would normally require considerable humility for a well-established professor to graciously accept devastating criticism of his work from one of his students. It may also require humility for an athlete to continue to acknowledge his weak points and his need for improvement—and perhaps the essential role of his coach—after a glorious Olympic victory.

Though at first blush the connection between humility and integrity may not seem obvious, when we consider what the whole-hearted pursuit of the good entails, the connection becomes more apparent: if I am to take appropriate steps to determine what is objectively good, and then to act on that determination—at least much of the time—then I will need to habitually resist the all-too-human tendency to exaggerate or over-estimate my own merit and strength; and I need to be willing to seek out the help and advice of people who are as wise or wiser than I. If a major aspect of doing the good is becoming a better person all-round, then I had better be disposed to make a realistic assessment of my character so that I can see where improvements are needed. Similarly, if discovery of the good is a social and interdependent affair in which “no man is an island,” then I had better be willing to acknowledge my dependency on others and rely upon them, especially in areas where my own capacities or experiences are more limited. Somebody who claims to be entirely self-sufficient and is unwilling to avail of the advice or assistance of others in the moral life strikes us as somebody who is either mentally incompetent (and thus not typically a candidate for attributions of integrity) or arrogant

---

and self-deceived, not genuinely committed to discerning and doing what is objectively good.

§4.5.3 Connecting the Form and Substance of Integrity

Many accounts of integrity insist that agents of integrity meet a certain threshold of “reasonableness” or decency in their actions. I believe that my account can offer a more cogent explanation of this intuitive link between form and substance than existing ones. On my account, there are at least two reasons we reserve the quality of integrity to those who meet a certain threshold of decency or goodness in their actions: first, because human beings, we believe, have similar cognitive tools or capacities for discerning the truth both in the theoretical and the practical sphere which, conscientiously employed, will be similarly responsive to certain aspects of reality. Consequently, a failure to see or do what is right or good may, in certain instances, betray a deeper failure to responsibly take appropriate steps to know the truth about one’s situation and make a good faith effort to act in accordance with that judgment. Secondly, as we have seen, integrity already requires a family of virtues or good-making dispositions, including temperance, conscientiousness, honesty, courage, and humility. The person who falls short on these virtues is to that extent lacking in integrity. In short, our shared cognitive capacities and the virtues that the conscientious pursuit of goodness entails place considerable normative constraints on what can count as a life of integrity. These constraints do not imply moral perfection, but they do provide a cogent rationale for

263 Here, I assume that values are part of reality, and I believe this is how they are treated when we undertake serious ethical and moral decisions. Beyond that, I am not proposing any special theory of values here.
deeming incompatible with integrity many forms of behavior such as cruelty, thoughtlessness towards others, and the rationalization of one’s vices.

Now, one obvious question raised by this brief argument is whether one can be so confident in people’s epistemic or cognitive capacities? For surely, we can easily imagine someone who inherits deep-seated prejudices and deeply offensive attitudes from their parents or culture and are, acting with full integrity (honestly, conscientiously, intending the good above all else, and so on), led to horrible moral conclusions, including the view that some human beings are inferior to others and do not deserve to be protected against violence? The first thing to say in this regard is that in a corrupt culture or with a corrupt upbringing, it may require extraordinary integrity to resist the force of habituation and social pressures of various sorts. However, I think dissenting subcultures in, say, Nazi Germany and antebellum south, demonstrate that our powers of ethical cognition are not entirely overwhelmed by a corrupt culture. Secondly, I grant that we will be harsher upon those who had a good formation in the virtues if they undertake certain evil actions. They “should have known better,” we will say, because they had the moral resources (viz., the virtues) required to resist the pressure of corruption. In other words, our judgments about people’s integrity may vary in degrees across different social and cultural milieus. Finally, I would suggest that certain forms of social and cultural corruption may even rob one of the tools necessary to exercise integrity, in which case a person’s integrity has been severely compromised before they even act. For example, Nazi culture hardly promoted the virtues of courage, honesty, conscientiousness, and humility in German citizens (although it may have promoted some or all of these virtues with respect to certain partial ends, e.g. courage with respect to war, but not with respect to seeking the good

264 Thanks to Chris Eberle for bringing this question to my attention.
at great personal cost)! To the extent that it undercut one or more of these virtues, it undermined people’s capacity to live lives of integrity.

§4.5.4 Integrity as a Matter of Degrees

There is at least one respect in which ordinary usage is misleading and unduly constraining: typically, we talk as if integrity were a simple property that an agent either possesses or lacks. We seemingly qualify for the attribution of integrity by meeting a certain threshold of decency or reasonableness. This threshold conception is reified in many recent theories of integrity, which associate integrity with a basic level of decency or reasonableness. But there is no principled reason for restricting the concept in this way, and we can find support in certain uses of the concept for treating integrity as a matter of degrees rather than as a matter of a threshold. What I want to suggest is that though we can accommodate the threshold sense of integrity, we need to move beyond it to a more open-ended sense.

It is helpful to consider the threshold sense of integrity as running parallel to the threshold sense of, say, the virtue of honesty: clearly, we might ask whether or not someone is honest, meaning, “is so-and-so basically honest?” or better, “do we have any special reason to believe that person is not to be trusted?” But we might also suggest, without any incoherence, that while John and Peter are both honest, John is more honest than Peter, or John displayed an extraordinary degree of honesty in disclosing self-

265 The “reasonableness” of principles plays a central role in the account of integrity offered in Cox, Caze, and Levine, *Integrity and the Fragile Self*, as it does in McFall, "Integrity," which argues that “when we grant integrity to a person, we need not approve of his or her principles or commitments, but we must at least recognize them as ones a reasonable person might take to be of great importance and ones that a reasonable person might be tempted to sacrifice to some lesser yet still recognizable goods” (11).
incriminating information that he was not required to disclose. There are “ordinary” or “basic” levels of honesty, and then there are “extraordinary” or perhaps, heroic levels of honesty. In a similar way, when we say, “does Elizabeth have integrity?” what we really mean is, “do we have any special reason to doubt Elizabeth’s integrity?” or, “does she reach a basic level of integrity, as far as we can tell from the evidence?” And yet, we can still maintain, with perfect coherence, that Elizabeth and Thomas More, while both possessing integrity, possess it in different degrees. Elizabeth is basically a woman of integrity: we have no reason to doubt her honesty, conscientiousness, etc. But Thomas More was by all accounts a man of outstanding integrity: he was put on trial for his responsibly-formed and pondered moral convictions and remained faithful to them even in the face of death. This example shows that integrity is a property we attribute to persons in degrees.  

§4.6 Objections and Replies

So far, I have developed in brief outline the essential features of my theory of integrity. But these features will need to be fleshed out further in response to a series of objections that could be leveled against my view. I will consider four such objections here: (1) Although my account requires that agents of integrity make a conscientious effort to discern what is objectively good, in practice the person of integrity can only act on his or her best lights, on his or her subjective grasp of the good. People often diverge dramatically from each other in their subjective grasp of the good, so the “virtue” of

Of course, to say of the inveterate killer that he has a “very small” degree of integrity would sound strange, since integrity is generally a term of praise, like virtue. Yet no stranger than saying that the thoroughly vicious person has a small degree of virtue. In both cases, the statements are probably true, unless we imagine someone who hasn’t a trace of goodness in him at all.
integrity can lead to social impasse and irresolvable conflicts. Consequently, the virtue of integrity cannot be valued or pursued unreservedly—it must be constrained both (a) *externally* by the State, whose concerns extend beyond the preservation of individuals’ integrity to the preservation of justice and the common good, and (b) *internally* by the individual, who must be concerned not only with his own integrity (this would be excessively self-indulgent) but with exercising social and political virtues such as justice and toleration in his dealings with other citizens. (2) This concept of integrity is so wide that it ends up dissolving into a morally good life, and thus loses its specificity. For surely, the dispositions required in order to have a stable and effective desire to do the good ultimately encompass all of the virtues? (3) Modern individuals find themselves thrust into a complex “web of group affiliations,” i.e. they belong to a multiplicity of social groups, each with its own set of demands and ethical constraints. In other words, we occupy different, and often conflicting roles in life. Full-blooded integrity can only be purchased in the modern world by retreating into social enclaves where roles are harmonious and well-defined. But this is unrealistic and morally questionable, given the benefits and freedoms associated with modern economic and social life, which cannot thrive without a proliferation of dynamic roles. (4) Finally, the ideal of integrity as perfect coherence is a misleading and dangerous one, since often the challenge of living well is not to *eliminate* conflict—which could be either fanatical or simplistic—but to manage it and where possible mitigate it. These, then, are four of the more interesting objections that can be brought against my account. As I answer each objection in turn, I hope that my account will be enriched and refined further, and the most important misgivings of the reader will be assuaged.
Objection 1: The “virtue” of integrity must be constrained both externally by the
requirements of justice and the common good duly interpreted by the State, and internally
by social virtues such as justice and toleration, otherwise it may be socially destructive.

Although my account posits a fairly strong connection between the virtue of
integrity and objective moral goodness, experience teaches us that people often diverge in
their subjective grasp of the good. Whether or not there is an “objective good” “out
there,” agents only have their own subjective perceptions to fall back on. Therefore, the
virtue of integrity, if pursued wholeheartedly, may often result in interminable
disagreement and even, in some instances, social anarchy. The solution to this problem
would appear to be to constrain integrity from without through the application of social
and legal sanctions when matters of public order and justice conflict with individual
integrity; and from within (that is, from within the actor himself), with other virtues such
as toleration and justice, which are more oriented towards the demands of social life than
integrity and less driven by a concern for personal well-being or righteousness.

Let us take the second part of this objection first, the part that speaks directly to
the agent: integrity should not be pursued wholeheartedly; instead, it should be pursued in
moderation, weighing other virtues such as justice against it. If this objection is right,
then integrity is just one property of agency, a single set of dispositions if you will, that
may have to be compromised when it interferes with, or conflicts with, other virtues or
goods. The first thing to be said in response to this worry is that the fact that somebody is
doing what she takes to be right or good, in and of itself, is not sufficient to warrant the
judgment that she is acting with integrity. It may well be, for example, that she is deeply
deluded about her own motives, or has not sought advice where she could and should have, or has disregarded important evidence bearing on her decision, or is driven to her conclusion by an excessive desire for pleasure.

Secondly, integrity is not just one more virtue concerned with some special sort of good, e.g. personal goods. Rather, it is a complex of virtuous dispositions whose object is the good, in whatever shape or form it comes. Therefore, to say to somebody that he needs to compromise his integrity in the interests of his own considered views of justice or toleration must presuppose that the discourse of justice and rights takes place outside or beyond the discourse of goodness, whether it be a matter of good states-of-affairs, good acts, or a good life. But to say to an agent, “Now, you are acting well/doing good, but at the cost of justice,” or “Now, you are acting justly, but at the cost of goodness,” is ultimately incoherent, unless good is defined so narrowly that it stands apart from what is just or right. Therefore, to the extent that the good life, i.e. the life well-lived, is obviously the just life, and the life that is sensitive to communal and social goods and demands, the person of integrity cannot coherently pursue personal goals without taking due cognizance of the demands of the community.267

Of course, there is no guarantee that the demands of the community will, as a matter of fact, line up perfectly with the judgments of individual members of the

---

267 My concept of integrity does not accept any radical distinction between a personal and impartial standpoint in morality (For one influential defense of an impartial perspective in ethics, see Thomas Nagel, *The View from Nowhere* (Oxford: Oxford University Press, 1986)). All moral points of view are the points of view of persons, who are personally committed to doing what is good and to becoming better people. In seeking to do what is good, nobody can completely abstract himself or herself from the decision-making process. Moral decisions always involves decisions as to which value(s) matter to me, which values I am to pursue, even if my immediate objective is to do good to another or to change the world for the better. When confronted with the need for social cooperation, the person of integrity must consider how his desire for, or evaluation of, cooperation, fits within his wider web of desires or evaluations. There is no simple trade-off between my “personal” conception of the good and an “impartial” criterion of value such as justice. On the contrary, my “personal” conception of the good contains within it some ordering, however tentative, of the various dimensions of goodness, including justice.
community concerning the good to be done and the evil to be avoided. Some level of conflict between the State and its laws on the one hand, and the ethical stances of individual citizens on the other, seems inevitable. But this is not a problem unique to a robust ideal of integrity: any theory of justice or political morality which wishes to preserve the moral autonomy and thoughtfulness of citizens must contemplate the possibility that citizens may conscientiously dissent from State decisions. On the other hand, to the extent that the good for humans is manifestly social and political and to the extent that interdependency is an essential feature of the human condition, no sane person of integrity, conscientiously discerning the good and ordering his life accordingly, could lightly set aside the demands of political and social order. Furthermore, citizens are accountable to others—in the first place, to their fellow citizens—for their ethical and moral decisions, at least those that directly affect the fate of their neighbors and the wider community. To say, “I acted with integrity” is not a justification of one’s actions: to justify one’s actions is to reveal the reasons that motivated them—and if one possesses a substantial degree of integrity, one has the humility and honesty required to acknowledge one’s errors when they become apparent, and the courage and temperance required to alter one’s behavior accordingly.

Third, the general suspicion of the responsiveness of integrity to social or moral demands may derive from the thought that, at least when it comes to our personal views of value or worth (as opposed to, say, our views of justice), we effectively inhabit separate moral universes. But this seems implausible to me. Granted, in many

---

268 Thus, it would seem that one of the basic assumptions supporting Rawls’s belief in the priority of the right over the good is the thought that matters of justice or right are generally more accessible to our shared reason than matters concerning the good, such as the nature of the good life and virtue, man’s ends, and so on. Yet the kinds of apparently interminable debates that go on in Western countries about the
important respects we may disagree on what is a good life and what is a bad or depraved life. Nevertheless, it does not follow that the concept of good presupposed by the complex disposition of integrity is a purely formal one: it is a rich concept, about which we know a lot already. We know what count as central and virtually undisputed cases of goodness, and what count as central and virtually undisputed cases of badness, and we hold people accountable for their ethical and moral judgments accordingly. For example, anyone who judges that torturing children for pleasure or imprisoning the innocent is good is either crazy, and thus not party to a serious moral conversation, or severely lacking one or more of the virtues of integrity, whether it be temperance, conscientiousness, honesty, prudence, courage, or humility. It is very hard to believe that a (even minimally) conscientious, temperate, honest, prudent, courageous, and humble inquirer, whose dominant and normally effective goal is to do what is objectively good, could reach a conclusion that is blatantly offensive to most minimally virtuous inquirers.

One of the sources of individualistic (and hence, anti-social) interpretations of the virtue of integrity is the distinctively modern assumption that the good can be pursued effectively in isolation from the wider community. Yet if we consider the conditions under which the good can realistically be grasped and attained, it becomes clear that people are radically interdependent in the pursuit of goodness.\textsuperscript{269} Right from infancy, a person’s very concept of goodness is taught through language and by learning from those around her. The pursuit of the good is deeply rooted in intersubjective or social practices in two ways in particular: first, our knowledge of the good, however inchoate, is learnt in

\textsuperscript{269} This is one of the principal claims defended in MacIntyre, \textit{Dependent Rational Animals}. 

\hspace{1cm}
large part through the example, teaching, and friendship of others, not only as children, but as adults. To take just one example, think about how we confer with people we trust about difficult decisions in which the best course of action is not evident to us. Also, consider how the disapproval of those we trust and admire can move us to reconsider a course of action we had previously deemed unproblematic. Secondly, we depend on each other to do and attain the good. This is true in many different ways. One obvious example is the good of friendship, which clearly can only be achieved with the willing collaboration of another. Another example is the social good of harmony or peace. This is clearly dependent on the goodwill and cooperation of our fellow citizens and even of non-citizens. Even goods that primarily concern the individual person, such as intellectual acumen, can only be attained through the example and instruction of another. More generally, a person is sustained and renewed in his struggle to live a good life by the affection, love, and support, of family and friends. Absent these, it is hard to imagine how we could persevere in doing what we believe to be right in difficult or trying circumstances. It is clear upon reflection that anyone who thinks he can either know or do the good independently of other human beings is displaying an inordinate and unfounded trust in his own epistemological and moral resources.

Now, consider the example of a corporate accountant whose decision to “cook the books” costs thousands of shareholders their life savings. When confronted with the evil of his actions, he throws his hands in the air and says, “What business of mine is their misfortune? I have a family to take care of.” Even assuming that he is speaking sincerely, and this genuinely reflects his considered opinion on this matter, we are unlikely to think of him as a person of integrity. Why not? There are probably many different reasons, but
at least one of them is the scant regard he pays to the social dimension of well-being, the fact that we are thrown into society whether we like it or not. We know from infancy that man is a social being, and that whatever his good is, it is inextricably entangled with the good of his fellows and of his society, even if this knowledge often remains an inarticulate sense of interdependency. If somebody acts blind to fundamental and obvious dimensions of the good, e.g. the social and political dimension, we will not just shrug our shoulders and put it down to the person’s idiosyncratic grasp of the good: we will seek for an explanation in terms of the vices that might undermine the quality or seriousness of his quest for the good. So somebody who turns his or her back on the requirements of life in society, or disregards the interests of his family or fellow citizens, is lacking integrity to the extent that he is overlooking an obvious dimension of the good—a dimension that should be obvious to any person who is exercising the virtues central to integrity—prudence, conscientiousness, honesty, courage, humility, and temperance— in the service of an effective and dominant desire to do what is good.

But perhaps this answer will not yet satisfy the reader: for surely, we should not rule out in advance the possibility that somebody may conceive her good in a fundamentally unjust way, in a way that is inimical to the interests of her fellow citizens or the interests of her society as a whole? Even granting that a person cannot coherently disregard the interests of others and claim to be living a life of “integrity,” some persons whose integrity we have no good reason to question, may interpret the interests of others in ways that are either obviously wrong, or ultimately detrimental to the common good. What is a suitable political response to this fallout of integrity? Surely the proponent of integrity will not suggest that we owe full respect and toleration, politically, to every
judgment and act that proceeds, on its face, from the virtues of integrity? Surely, we cannot accept that the right to act with integrity should serve as a trump card against any and all State interventions on behalf of the common good, the interests of citizens, etc.? Integrity, even if it is a laudable trait, needs to be externally controlled or confined by the State whenever it threatens to undermine public order, liberty, equality, or any of the values characteristically protected by the liberal state.

At no point do I want to claim that a right to act with integrity serves as an automatic trump card against the right of the State or government to regulate public life in accordance with justice and the common good. Nor am I claiming that a life of integrity should be unrestrained by the expectations and demands of one’s society. After all, the learning of social norms is a central aspect of our moral education. However, I do want to argue that in general, integrity is a politically and socially salutary set of virtues; and that attempts to systematically curb or confine it for political reasons, whether legally or socially, are deeply misguided. Integrity is politically and socially salutary, because it makes citizens better-disposed to discover the truth and take on board the full complexity of social and political problems in their deliberations. This runs contrary to certain stereotypes which associate integrity with politically destructive traits such as rigidity, self-indulgent moralism, inability to compromise, and simplistic or reductive outlooks on life. But the person of integrity is the person whose desire to do what is good is more important than any desire to be justified or come out “looking good”; the person who is conscientious in her quest for truth and thus loathe to settle for a simplistic or poorly thought through view of things. Finally, the person of integrity is someone who is willing to make political compromises where he conscientiously judges that such compromises
are either necessary or desirable for the sake of political order and the interests of others. But he is also someone who responsibly sets the limits of compromise in accordance with (a) his best judgment about acts and states-of-affairs that are unacceptably harmful or vicious; and (b) his best judgment about the prospects of achieving a superior outcome by refusing to compromise, naturally balancing the possible social and political costs and benefits of this refusal against the social and political costs and benefits of compromise.

For all that, I am not claiming that acts of integrity are infallible or that people of integrity get things right all the time. I am claiming, rather, that there is no special reason to think integrity is any more politically destructive than honesty, or prudence, or courage. And insofar as we do regulate the lives of persons of integrity, whether legally or through social norms, it should not be because of some structural feature of the lives of persons of integrity, e.g. their ongoing attempts to bring all of their values and commitments into harmony with each other; but because of the inherent demerit, evil, or injustice, of their attitudes, dispositions, or actions in particular cases.

In short, I am not making an argument here for the open-ended toleration of social and political acts springing from integrity: but I am arguing that we must reject any argument, such as that we find in political liberalism, that systematically delegitimates efforts to live with integrity without considering the merits of the actions, attitudes, or dispositions at issue. There is nothing inherently suspect or illiberal about speaking and acting in light of the full array of one’s beliefs, desires, and values, ordered according to one’s grasp of the good. If we want to question somebody’s seriousness or commitment to democracy, it is not enough to say, “look, you’re just acting from your deepest personal beliefs and values, and refusing to show how your actions can be justified by
public political values.” Rather, we have to say, “these beliefs and values you are invoking now are inappropriate grounds for this sort of policy,” because they are either false, irrational, narrow-minded, harmful, or peculiarly inaccessible to the people involved. Essentially, I am arguing that a general moral suspicion of integrity such as that implied by Rawls’s duty of civility cannot be sustained, while individual acts springing from (some degree of) integrity may be criticizable or even deemed illegal for a variety of reasons not derivable from a general suspicion of integrity.

**Objection 2:** *This concept of integrity is so wide that it ends up dissolving into a morally good life, and thus loses its specificity.*

The second objection against my account goes something like this: if you equate integrity with a family of dispositions required in order to (a) have a highest-order desire to do what is good, (b) take appropriate steps to determine what is good, and (c) order one’s moral life accordingly, you end up not with a nuclear family of dispositions, but a disturbingly extended family of dispositions, perhaps even all of the virtues. For what virtue would we say is *unnecessary* in order to “take appropriate steps to determine what is good, and order one’s life in accordance with that determination”?

The first thing to say in response to this objection is that my account does not imply any strict identity between integrity and goodness. It allows that a person who orders his life in accordance with his responsible discernment of the good may be simply mistaken about what the good requires. Thus, even though he is exercising the virtues to some degree (e.g. honesty, courage, prudence), he is clearly not perfectly virtuous if the object of his life—his view of the good—is somehow deficient. In short, there is no
simple equation of moral goodness and integrity on my account. Integrity is a necessary condition for full moral goodness, but not a sufficient condition.

Secondly, it should not surprise us that in order to integrate or order one’s desires, values, belief, and actions, and dispositions, both among each other and in accordance with a serious discernment of the good—even if that integration is always a work in progress—we will have need of many good character traits. The ones I have discussed—temperance, prudence, courage, honesty, humility, conscientiousness—are not meant to constitute an exhaustive list, but I do consider them to be especially fundamental to the life of integrity, for reasons I have already suggested. And the extensiveness of this list does not destroy the specificity of integrity, since when we say somebody has integrity we are not just referring to a random list of virtues, but to the kind of character a person must have if he is to order his beliefs, desires, evaluations, and actions, over time, so that they cohere with each other and with the good as he discerns it. So integrity is not just a label for the virtues, for which we could substitute, say, the word, “virtue,” but a property of character that illuminates the virtues from a particular perspective. Because integrity does not refer to any discrete disposition but to a complex character type, indeed to the whole moral life looked at as a continual quest for the good, any plausible definition of integrity is likely to imply the presence of many of the virtues. But it does not, on that account, fail to offer an informative and insightful interpretation of the moral life.

270 Indeed, in order to undertake any discrete virtuous act, it should not surprise us that we will require, to some degree, many if not all of the other virtues. It is hard to imagine, for example, a person undertaking a genuinely compassionate act without also possessing the virtues of justice, prudence, and courage, at least to some degree.
**Objection 3:** The modern proliferation of social roles forces us to choose between integrity and modern life with its “complex web of group affiliations” and proliferation of roles. Surely we cannot afford to retreat from the complexity of modern life?

One of the most compelling objections against integrity as a desirable quality is the complexity of modern life, which at times appears to not only sanction, but require, a compartmentalization of one’s life in different, and often only tenuously connected, roles. This objection suggests that distinct social roles are so heterogeneous and conflicted in their demands that to impose a requirement of integrity is essentially to ask people to turn their back on modern life, to retreat into some social enclave where roles are coherent, few, and well-defined. One of the most popular examples in this regard is the role of politician or statesman, which often requires, it is said, immoral or morally questionable behavior, such as lying, in order to further the country’s political interests. The politician can perhaps afford to avoid lies in his private life, but when a country’s interests are at stake, he cannot afford the “luxury” of such a sensitive conscience.

The best way to address the concern about role proliferation and its impact on the unity of a life is to start with the concept of a role, and the ways a role can place ethical demands or constraints on a person. Drawing on the definition of role offered by the OED\textsuperscript{272}, I would suggest the following rough definition of social role: A social role is (i)

\textsuperscript{271} Alasdair MacIntyre (disapprovingly) highlights the modern tendency towards compartmentalization of roles in *After Virtue*: “the unity of a human life becomes invisible to us when a sharp separation is made either between the individual and the roles that he or she plays…or between the different role- and quasi-role-enactments of an individual life so that life comes to appear as nothing but a series of unconnected episodes—a liquidation of the self…” (204).

\textsuperscript{272} “1a. The part or character which one has to play, undertakes, or assumes. Freq. fig., with reference to the part played by a person in society or life. 1b. The typical or characteristic function performed by someone or something; freq. in phr. to play a role. 2. Social Psychol. The behaviour that an individual feels
a *function* within society constituted by (ii) one’s social position or relation towards others and (iii) social and institutional norms, goals, and standards of excellence associated with the function, which (iv) shape the practical reasoning of the role-holder and related parties in distinctive ways. Unpacking this definition further, we can say that insofar as one occupies a social role, one is expected to, and usually required to, perform a set of functions constituted by certain relations to others and to society. Such functions are characteristically derived from relations with others via special responsibilities towards others, shared interests or goals, or some special affection or love for another. For example, the role of teacher involves the *function* of educating children (or adults) which applies to someone just insofar as he or she is situated vis-à-vis students in such a way as to contract a *responsibility* to educate them. This may be because he has assumed the paid post of teacher, or it may be because there is no one else available to educate a child under his care. The important point is that the *function* (or functions) entailed by the role of teacher arises out of a special relationship with and associated responsibility towards a finite group of persons.\(^{273}\) The nature of this relationship and its associated responsibilities are defined by general social norms, expectations and goals attached to the relationship, to be interpreted by the role-holder and related parties (in this case, most obviously, the students).

Part (iv) of the definition of role makes the important point that the function associated with a role is not like the function of, say, an organ, that determines behavior it appropriate to assume in adapting to any form of social interaction; the behaviour considered appropriate to the interaction demanded by a particular kind of work or social position.\(^{273}\)

\(^{273}\) There is an interesting literature exploring the notion of role obligations, or obligations rooted in special relationships, and how *special* obligations can (or cannot) be reconciled with universal Kantian obligations. See, for example, Michael Hardimon, "Role Obligations," *The Journal of Philosophy* 91, no. 7 (1994): 333-363.
mechanistically; rather, it is an important input into the role-holder’s and related parties’ deliberations and decisions. The function of policeman, or prime minister, or teacher, or citizen, does not immediately compel any particular form of behavior: instead, it shapes the reasoning and decisions of the role-holder and related parties (e.g. beneficiaries or patrons of the role). This is true in two senses: first, in the obvious sense that a person is free at any moment to abandon, disregard, or sabotage a function, which suggests that the function does not directly determine behavior or decisions. But it is also true in the less obvious and more interesting sense that given the complexity of human life and human goods, any given function one performs in society can hardly provide a definitive or exhaustive set of considerations bearing on action: one’s role or function provides one set of practical (that is, action-directing) considerations among others (including one’s other roles, one’s humanity, and one’s general beliefs about value). I do not mean by this to downplay the centrality of one’s role to deliberation in certain role-related contexts, but I do mean to emphasize that one’s role can never be the exclusive and overriding basis of one’s practical reasoning.

Now, consider how a proliferation of roles might undermine the integrity of an agent: in the first place, different roles might impose conflicting responsibilities or tasks upon an agent, either in the straightforward sense that the respective tasks/responsibilities cannot all be feasibly realized (for example, a mother who is also a CEO may find it impossible to realize all of the characteristic responsibilities of the two roles), or in the deeper sense that the different roles reflect mutually irreconcilable values. In the latter case, even if both roles can be feasibly implemented, they generate a conflict among the agent’s evaluations and corresponding actions.
Secondly, some roles may impose upon an agent responsibilities or tasks that reflect values that conflict with the agent’s own evaluations and desires as a person. Here, we are not dealing with conflicting roles, but with a conflict between a particular role and the values a person cherishes independently of any specific role. To be a person or moral agent is not to fill a role, but to exist as a certain sort of creature, capable of intelligent and free acts. That personhood or agency is not a role becomes clear if we recall our definition of role: “a \textit{function} within society constituted by one’s social position or relations towards others, which shapes the decisions of the role-holder and related parties in distinctive ways.” Personhood or agency is not constituted by a social function, but by the capacity for reflexive deliberation and choice. Of course, a person may undertake a series of functions, but there is no distinctive function which marks out the identity of a person and can be distinguished from other functions.\footnote{I do not mean by this to deny (or affirm) that persons have a \textit{natural} function just by virtue of the kind of being they are. But I do mean to deny that their identity as persons is definitively captured by the \textit{social} functions they perform or their social relations. No doubt a person’s role in society says a lot about who he or she is, but this is certainly not what we mean when we say so-and-so is a \textit{person}.}

There is no disputing the fact that role differentiation and roles in general can threaten an agent’s integrity as I have defined it. Consequently, anyone who is serious about living a (maximally) integrated life must take every reasonable step to negotiate and eliminate conflicts rooted in diverse roles he occupies. This is a delicate and often trying task. In many cases, it may be that one simply \textit{cannot} achieve full integration both among one’s diverse roles and between one’s roles and one’s personal beliefs and values. In that case, the agent is not called to abandon the goal of integrity but to accept that for now at least, he cannot achieve it to the extent he would like. There are also different strategies he can use to deal with role-based threats to integrity: one is to step...
back and critically evaluate the role, to see whether one can implement it selectively in a way that preserves one’s integrity. Another is to decide that the role is so comprehensively at odds with one’s integrity, or the circumstances militate so strongly against compromising one’s role performance, that the threat to integrity can only be adequately confronted by abandoning the role entirely.

If an effort to live with full integrity resulted in the systematic dismantlement of most modern roles, then we would have good cause to reassess the goal of integrity and inquire whether it is in fact quite utopian in the context of modernity. However, I do not believe that such a result is likely, because I believe the radical compartmentalization of roles is an optional way of thinking and acting, not an inevitable outcome of modern life. In the case where a single person occupies multiple roles, i.e. performs multiple functions in a society, and thereby contracts multiple responsibilities towards different groups, it is the special responsibility of that person to make every effort to ensure that the demands of his diverse roles do not conflict, either mutually, or with his own role-independent values, desires, and beliefs. Where a conflict does arise, the agent needs to assess his prospects for creatively re-shaping one of the roles without violating the responsibilities and commitments associated with the role; or revising his existing habits of action, evaluations, desires, or beliefs to accommodate the role. If neither of these options can be reconciled with a fundamental commitment to the good (the highest desire of the person of integrity), an agent may feel compelled to either implement the role selectively or refuse the role entirely.

What gives the case against integrity from role proliferation more force than it deserves is the false notion that roles are rigidly defined functions that are adopted by
agents on a “take it or leave it” basis. Although some roles are more narrowly defined than others, the definition of most roles is sufficiently vague and open-textured that they are susceptible to re-interpretation and modification by creative and thoughtful persons. For example, the role of waiter is primarily to serve customers their meals. However, someone may choose to work longer or shorter hours, and enter more or less wholeheartedly into his role of serving others. He may at first find the role demeaning and inhuman, but gradually learn that some friendly banter with his customers and seeing them as fellow persons instead of cash dispensers, elevates the role in his eyes and reconciles it with at least some of his important aspirations or desires.

Now, take a more problematic case: the case of the member of Congress who comes from a predominantly pro-choice district but believes abortion amounts to the unjust taking of innocent lives. How can he represent the views of his constituents on this delicate issue, without compromising his personal integrity? Either he can just bite the bullet and go against the wishes of his constituents, in which case some would say he is acting without a democratic mandate, or he can publicly support, or at least not oppose, the views of his constituents on this issue, in which case some would say his is acting contrary to his deepest values, and thus lacking in integrity.

The dilemma painted here is real, but misleading. Members of Congress are clearly expected to fulfill their function of political representation in a way that is sensitive to the views of their constituents. However, it is widely accepted that the function of political representative encompasses the role of delegate as well as trustee.\footnote{For an incisive and well-rounded account of the various dimensions of representation, see Hanna Fenichel Pitkin, \textit{The Concept of Representation} (Berkeley: University of California Press, 1972).} The delegate is meant to simply act as his constituents want him to act, insofar as that
can be ascertained. The trustee, on the other hand, is meant to act in the interests of his constituents, making an all-things-considered judgment that surpasses and occasionally contradicts their opinions. The delegate and trustee dimensions of representation illustrate well the vagueness and open texture of social roles, which call for prudent and dynamic interpretation rather than unthinking acquiescence. To return to the pro-life representative: if he believes, in a reflective rather than unthinking way, that abortion is a wrong serious enough to outweigh the contrary opinions of his constituents, then to publicly oppose abortion would seem to be required by integrity. Whether it is ultimately in conformity with integrity depends, of course, on whether the representative’s actions are truly animated by the virtues of integrity, most centrally honesty, temperance, courage, conscientiousness, and humility. But the representative, like any human being, can do no better than act on his best judgment, conscientiously relying on the moral and epistemic resources available to him. To act on one’s best judgment where that judgment is conscientiously and responsibly formed is a necessary and central condition for the attainment of integrity, and therefore tends to win people’s admiration. It remains possible, of course, that the representative, his integrity notwithstanding, is ultimately misguided about the good.

Now, take the example given by Rawls as a sort of exemplar of public reason, the role of a judge, and imagine a judge who finds his role conflicts with his ethical convictions. Consider first what is distinctive about a judge’s function: a judge’s function is to interpret and apply the law, and this sets him apart from other political actors such as legislators, police officers, and ordinary citizens. The primary role of a police officer is to enforce the law, though clearly this function entails an interpretation
and application of the law to the particular case he confronts. The police officer does not have the luxury of detachment and time for reflection combined with the training in legal interpretation that a judge’s function requires. Legislators have the function of designing new laws and policies and reforming existing ones to the benefit of the wider community. The legislative function differs significantly from the judge’s in that legislators can (a) invent laws ab initio, rather than merely interpret and apply existing laws; (b) act from their own initiative, or at will, without any particular case being brought before them. Finally, citizens are required by their role to contribute to the common good of their political community in whatever ways their circumstances and prudence permit, including obeying the laws of their country (except in extreme cases of injustice) and supporting laws and policies according to their best judgment about what would benefit the community. Citizens thus have a very general function that gets expressed differently according to the more specific role a citizen assumes within the political community (e.g. police officer, voter, legislator, judge).

Now, imagine a case where a judge finds that his own personal ethical convictions about justice conflict with what he takes to be the most accurate interpretation of the law. What should he do? Isn’t this a case where he needs to set aside the impetus towards integrity and bow to the requirements of his role? I would argue that on the contrary, this is a case where the person of integrity must decide whether continuing to occupy the role of judge is compatible with his integrity. For example, it may be that the injustice of the law is not so severe that applying an imperfect law is wrong or unethical. Or it may be that the law is so egregiously unjust (e.g. a law treating black men and women as property) that the judge finds himself compelled to resign
rather than apply it. Or perhaps the judge can resign and dissent publicly, at least in extreme cases where he feels the fundamental spirit and purpose of law has been nullified. The point is that each case must be considered in its proper context, which requires both an inquiry by the judge into what is required of him as a virtuous person; and an inquiry into what is required of him by his role as judge and as citizen.

Finally, Rawls might argue that just as a judge is circumscribed by his role and its requirements (viz., interpretation and application of the law), so a citizen is circumscribed by his role and its requirements, including the duty of civility or the duty to offer publicly accessible reasons for one’s political advocacy. Building on the example of a judge, Rawls might argue that a citizen who rejects the norms of public reason, or the duty of civility, may be acting on his best judgment, but he is repudiating the role of citizen and its requirements. He is now acting not in a public-spirited, but in a private-spirited way, engaging in sectarian arguments that little befit the public discourse of citizens.

A few points need to be made to counter this sort of argument: first, the duty of civility that lies at the heart of public reason is both unrealistic and incoherent, as I showed in chapter 2, and therefore it is not a plausible interpretation of public reason. Secondly, a citizen may make a good faith effort to render his arguments rationally accessible to others without acceding to Rawls’s stringent requirement of reciprocity: there is a large gap between Rawlsian publicity or reasonableness on the one hand, and completely sectarian, nonpublic argumentation on the other. For example, somebody may advance a natural law argument against same-sex marriage that fails Rawls’s stringent reciprocity requirement yet is, in some significant sense, rationally accessible.
to others. Thirdly, a judge has the limited function of interpreting and applying existing law, not inventing new laws, and it therefore makes a lot of sense to narrowly circumscribe his arguments within the legal domain. A citizen, on the other hand, is called upon to make broad and probing judgments about the common good, not just about existing laws, and these judgments, while they ought to be sensitive to existing customs and laws, are not strictly bound by them—a citizen may frequently advocate the repeal of ill-considered laws, for example. In other words citizens, at least acting in their ordinary, general capacity, are not just interpreters of law; they are makers of law. Consequently, they must invoke much more far-ranging and open-ended considerations than judges as they exercise their role.

What I hope to have shown in response to the worry about role differentiation is that although roles may and often do give rise to tensions within a person’s beliefs, evaluations, desires, and actions, these tensions need to be negotiated and eliminated to the extent practicable rather than settled for. Coping with these tensions may occasionally involve the partial or complete abandonment or rejection of certain roles, but roles are often sufficiently vague and open-textured in their definition that agents may dynamically adapt themselves to their roles, both revising their own beliefs, values, and desires in response to the role, and re-interpreting the role in light of their existing beliefs, values, and desires.
Objection 4: Often, living well requires us to acknowledge and struggle with moral conflict rather than eliminate it. Therefore, an ideal that emphasizes perfect coherence among commitments breeds fanaticism and self-deception.

If the fully integrated life is a life in which our desires, beliefs, actions, evaluations, and dispositions form a seamless whole, a life from which all traces of internal strife have been eliminated, then one might argue that the fully integrated life is not an appropriate goal for human agents to aspire to. After all, conflict and indecision often must be confronted and managed honestly rather than eliminated for the sake of integrity. Only the self-deceived and perhaps the arrogant person could take steps to eliminate all conflict from his life and then believe that he had succeeded. Surely wholeness of life literally speaking could only be grasped at by either suppressing significant parts of the self (e.g. deep desires) or simply through some form of self-deception. This is not to deny that some degree of coherence is an appropriate and desirable goal for moral agents to aim at. But it is to say that integrity is more about accepting and managing with honesty the fact of conflict than about achieving a coherent self.\footnote{This objection against integrity as wholeness finds its most cogent articulation in Cox, Caze, and Levine, Integrity and the Fragile Self, esp. chap. 2, “Integrity as a Virtue”: “As we conceive it, integrity is not a kind of wholeness, solidity of character or moral purity. It involves a capacity to respond to change in one’s values or circumstances, a kind of continual remaking of the self, as well as a capacity to balance competing commitments and values and to take responsibility for one’s work and thought. Understanding integrity involves taking the self to be always in process, rather than static and unchanging or containing an inner ‘core’ around which reasonably superficial changes are made” (41).}

There are certainly ways of pursuing coherence or self-integration that are unhealthy and even dangerous. However, to pursue the good under the aspect of wholeness and completeness, to have not just any desire but the desire for goodness as one’s controlling and integrating desire, prevents the sorts of pathologies that might be
encouraged by more formal ideals of integrity. On my account, the person of integrity endeavors to integrate his beliefs, actions, desires, evaluations and dispositions into a coherent whole. However, the goal of integration cannot be divorced from the goal of attaining the good, or if it is, then integration becomes morally vacuous. Thus, if a person’s grasp of the good remains unclear or divided, it is not a matter of integrity for the person to forcibly integrate his beliefs, desires, actions, evaluations, and dispositions according to some design that is arbitrary with respect to the good (as he perceives it). On the contrary, it is a matter of integrity to acknowledge the conflict and only resolve it insofar as he sees that it can be genuinely resolved in conformity with the good. Thus, the person who seeks an artificial or forced integration of his moral life, what some might call a “fanatic,” cannot plausibly be thought of as a person of integrity, for at least two reasons: first, integrity requires not only wholeness or oneness, but completeness, as I explained earlier. This means that the goal of integrity cannot justify aggressively seeking coherence at the cost of important aspects of the good life. The person of integrity not only seeks formal coherence among desires, but a certain completeness of desires; one’s desires ought to express at least the most important aspects of who one is, not just one aspect of one’s identity. Secondly, we would expect the person of integrity to take whatever reasonable steps are within his power to ensure that his knowledge of the good and his pursuit of the good are reliable. The responsible agent not only seeks for coherence for its own sake, but he seeks for coherence in the pursuit of good. Thus, he does not settle for whatever view of good he first comes across, but he makes a conscientious effort to determine what is objectively good for him (and for others). We could describe this as epistemic responsibility. It will not do to just latch onto some

---

277 See §4.1 above.
simplistic principle of integration and adopt it at a whim: we must struggle with our conflicts and use all the tools available to us, including our own native intelligence and skills and the assistance of those around us, to determine the true nature of the good and the limits as well as the possibilities of genuine integration at any moment.

***

I have undertaken this discussion of integrity with two basic goals in mind: first, to support the integrity-based case against Rawlsian public reason by spelling out in greater details what integrity amounts to and showing that some important liberal objections against it mischaracterize or misunderstand the notion they are attacking; and second, to advance our understanding of integrity and thus make a significant contribution towards recent debates among moral philosophers about the meaning of integrity.

It may be that, as this last objection suggests, conflict is a fact of life that will never go away. It certainly seems obvious that nobody in this life is a perfectly integrated agent, with all his or her desires, actions, and evaluations ordered according to a single unifying principle, the good. It also seems to be the case that anyone who *claimed* to have eliminated all conflict from their lives would be lying, whether to others or to themselves. This is all well and good. However, it does not in any way undermine the ideal of integrity, anymore than the temptation to think of oneself as perfect undermines the ideal of moral goodness. It is obvious, from the perspective of integrity, that we ought to be honest about the imperfections and inadequacies within our lives. Anything less would
constitute a form of self-deception or arrogance. However, nothing stops someone from leading her life guided by an ideal, such as integrity, which she has almost no hope of perfectly achieving. I doubt that anybody in this life ever achieves the fullness of integrity, but that is not the same thing as saying that people should stop pursuing integrity or measuring the success of their lives against the yardstick of full integrity.
CHAPTER 5: THE LIMITS OF CONTRACTUALIST ACCOUNTS OF CITIZENSHIP

In chapters 2 and 3, I offered an explication and critique of one of the most influential ideals of citizenship to be articulated by recent political philosophers, that advanced by John Rawls in Political Liberalism. I showed that Rawls’s ideal failed in several crucial respects: first, internally, by offering a test of political legitimacy (reciprocity) that his own theory could not come close to satisfying, and by proposing a normative basis for discourse that held out little or no realistic prospect of winning the stable allegiance of citizens; and second, externally, by offering an ideal of citizenship that gives short shrift to the virtue of integrity, and thus cannot be embraced wholeheartedly by a range of citizens otherwise sympathetic to liberal democracy. Since I identified political liberalism’s hostility to integrity as one of its salient defects, something more constructive needed to be said about this virtue than I, or other critics of political liberalism, had said. This gap was filled in chapter four, where I built up a useable concept of integrity, with a view to exhibiting its conceptual contours and socio-political import, and meeting some of the more challenging socio-political objections to a demanding ideal of integrity.

In this fifth chapter, I present a deeper and broader diagnosis of the failure of political liberalism to convincingly address the problem of political order in modern societies, less from the point of view of political liberalism than from a broadly liberal
perspective, i.e. a perspective that views as good and legitimate in principle, a range of values, practices and institutions associated with liberal democratic societies, such as representative government, rule of law, freedom of religion and conscience, a free market economy, and so on. In addition, the wider practical and theoretical significance of my critique of Rawls will become clearer in this chapter as I survey two prominent rivals to political liberalism, deliberative democracy and agonistic pluralism, and suggest that partly because they share with Rawls the intellectual heritage of the social contract tradition, they fall prey to similar and equally serious limitations. An awareness of the limitations of contemporary contractualist accounts of political order suggests that we ought at least to consider the possibility that another long-standing and respectable tradition beyond social contract theory—the classical tradition of the virtues—might serve to reinvigorate and support the modern quest for a just and stable political order.²⁷⁸ Indeed, that is the hypothesis that I explore in chapter 6. I hope that this chapter’s examination of political liberalism and two of its main contractualist rivals will convince the reader to give that hypothesis serious consideration.

I will begin my broader diagnosis of the limitations of political liberalism by (§5.1) setting out in advance what I take to be some conditions any modern ideal of citizenship must meet if it is to be a politically and morally viable response to the problem of politics as we encounter it today in liberal democratic societies. Second, guided by these criteria, (§5.2) I will argue (partly drawing together what has gone

²⁷⁸ Interestingly, the order of this inquiry, beginning with anomalies that can be recognized as such from within the liberal contractualist tradition, and proceeding to consider the resources of other traditions, seems to be the kind of thing Alasdair MacIntyre has in mind in his conception of constructive dialogue between apparently incommensurable traditions. See Alasdair MacIntyre, *Three Rival Versions of Moral Enquiry: Encyclopedia, Genealogy, and Tradition* (Notre Dame, Indiana: University of Notre Dame Press, 1990).
before) for the inadequacy of political liberalism as a response to the problem of politics, and thus the timeliness of a radical alternative. Thirdly, (§5.3) I offer a still broader and more penetrating diagnosis of the failure of political liberalism, suggesting that its limitations are attributable, in part, to its embrace of a contractualist vision of society. To make this suggestion more plausible, I consider two prominent alternatives to political liberalism, both of which I consider to be under the sway of contractualist assumptions: Iris Young’s ideal of deliberative democracy on the one hand, and Chantal Mouffe’s ideal of “agonistic pluralism” on the other. I will argue that each ideal can, like Rawls’s, be described as contractualist rather than aretaic in its basic orientation, and that both accounts fail to meet my suggested criteria of success, in no small part due to their contractualist foundations.

§5.1 Some Criteria for a Successful Account of Citizenship

Many political theorists follow Rawls in understanding the problem of politics in contemporary liberal democratic societies to be expressed in one fundamental question: “how is it possible for there to exist over time a just and stable society of free and equal citizens, who remain profoundly divided by reasonable religious, philosophical, and moral doctrines?” My understanding of the problem and the parameters of a successful answer are not strictly Rawlsian but they are broadly liberal. By this, I mean a philosophical outlook that accepts in principle if not always in practice, the legitimacy and goodness of the fundamental institutions and practices characteristic of liberal

---

279 PL, 4.

280 For example, just because one accepts the legitimacy of liberal institutions does not mean that one always approves of the uses to which they are put.
regimes, such as private property and the free market economy, personal liberty within the limits of the common good, equality before the law, freedom of religion, freedom of conscience, freedom of association, rule of law, democratic procedures, and the right to due process.

Since this chapter (and indeed this dissertation in its most basic questions) is concerned with the problem of politics understood in broadly liberal terms rather than with the peculiarities of political liberalism, I prefer to restate the problem in more ecumenical terms, without relying on epistemologically and morally loaded terms such as “reasonable.” In that spirit, the fundamental problem of politics in modern liberal democracies can be restated as a slightly modified form of Rawls’s question: “how is it possible for there to exist over time a just and stable society of free and equal citizens, who are profoundly divided by religious, philosophical, and moral doctrines?” Any politically and morally viable account must constitute a plausible answer to this question, an answer that strikes some balance between despairing “realism” on the one hand and unrealistic utopianism on the other. This balance is captured in three central conditions any successful account must meet: first, it must offer a plausible understanding of respect for persons; second, it must illuminate conditions of political stability that can serve as a realistic regulative ideal; and third, it must accommodate the demands of integrity, not absolutely, but to the extent practicable.\(^{281}\)

Let me explain briefly what these conditions amount to. The first condition, the need for a plausible understanding of respect for persons, is vital since interpersonal

\(^{281}\) Of course, I am not suggesting that these three conditions exhaust the conditions a plausible modern political theory must meet, but they are sufficiently pivotal and challenging that they can serve as a useful benchmark for comparing political liberalism with my aretaic account of citizenship.
interaction is at the heart of citizenship, and the recurring question confronting citizens in their public life is less one of law (most citizens are law-abiding) and more one of ethos, specifically, “how do I best show respect for my fellow citizens?” Some might say that it is not a matter of respect but of justice: “how can I act with justice towards my fellow citizens?” I would contend that respect encompasses justice in the sense that justice is a necessary but insufficient condition for respect. Respect is a broader and richer concept than justice and thus is not intended to substitute but to enlarge the virtue of justice.

The second condition of a successful account of citizenship is that it must illuminate the conditions of political stability that can serve as a realistic regulative ideal. A political regime is stable if a large majority of its citizens are loyal to its institutions and view it as worthy of their allegiance; disaffection from the regime and plots to disrupt or overturn its institutions are relatively contained, have little public support, and have a very slim probability of success. A regime that is stable focuses most of its energy on quotidian policy questions rather than on solving crises of legitimacy or averting widespread political disaffection or political acts of violence. For example, Northern Ireland during the “Troubles” when its politicians could hardly speak to each other, let alone conduct the affairs of government together, was politically unstable as its political energy was focused on averting violence and addressing questions of legitimacy rather than solving more quodian problems of social and economic policy, “normal” crime prevention, and so on.

Political stability is a constant work in progress, never a definitive achievement. Furthermore, it is a matter of degrees rather than an on-off property. Regimes and societies can be more or less stable—an ideal of citizenship should illuminate the types of civic interaction that can enhance stability, and thus serve as a regulative ideal that
encourages stabilizing behavior. Like all regulative ideals, it will be aspirational, but it must propose models of behavior that citizens can realistically approximate. For example, an ideal that posited perfect harmony or philosophical agreement between citizens would be unrealistically utopian and completely beyond the capacities of humans. Such an ideal would thus fail to serve as a realistic regulative ideal for civic life.

The third condition of a successful account is that it must accommodate the demands of ethical integrity, not absolutely, but to the extent practicable. Recall that according to my account, a person of integrity approximates the following conditions (their full embodiment would constitute full integrity): (a) his highest-order desire is to do the good; (b) that desire is effective and stable enough to dynamically structure and restructure his moral life (beliefs, actions, evaluations, desires, dispositions) over time; and (c) his beliefs, actions, evaluations, desires, and dispositions are mutually coherent. Clearly, no ideal of citizenship can give citizens absolute freedom to act with integrity unimpeded by the law or by social customs and norms. However, ideals of citizenship ought to be sensitive to the demands of integrity at least to the extent practicable, and certainly not attack or undercut them without a compelling public purpose.

---

282 Of course, in certain circumstances, it may be desirable or even morally obligatory to provoke political instability, for example, if the political system is no longer serving the interests of citizens. But I am assuming that under normal circumstances, stability and peace are desirable from a broadly liberal perspective.


284 See §3.2.2 above for a detailed analysis of the concept of reasonableness in political liberalism.
To recap, three conditions any viable ideal of citizenship must meet are: (a) it must offer a plausible account of respect; (b) it must realistically illuminate the conditions of political stability; and (c) it must be duly sensitive to the demands of integrity, to the extent practicable.

§5.2 Why We Need a Radical Alternative to the Rawlsian Ideal of Citizenship

Rawls’s rationale for limiting an account of civic virtue to the cooperative virtues, or to the virtues associated with “reasonableness,” is that a more full-blooded account of civic virtue could not secure the rational allegiance of a preponderance of intelligent, cooperative (what Rawls would call “reasonable”) citizens in a morally and religiously pluralistic democracy. Rawls’s account of civic virtue is meant to provide a regulative ideal of citizenship that can mediate philosophical and religious differences, an ideal all well-disposed or reasonable citizens can aspire to irrespective of their philosophical or religious backgrounds. However, Rawls’s ideal fails to meet the three criteria I have outlined for a successful account: first, its conception of respect for persons is philosophically implausible and politically divisive; second, its understanding of the sources and remedies for political instability is highly constricted; and third, it is

285 As we have seen already (§3.2.2), the concept of reasonableness in political liberalism is notoriously difficult to pin down. One of the main difficulties with the concept is that as explicated by Rawls, it often seems to involve a commitment to controversial parts of Rawls’s theory such as the duty of civility. By that account of reasonableness, anyone who disagrees with the duty of civility is ipso facto “unreasonable.” Independently of Rawls’s position, I believe a more plausible rendering of the concept from a broadly liberal perspective is that “reasonable” persons are persons who are responsible, fair-minded, cooperative, and rational (see §3.3.2). This leaves room for “reasonable” persons to argue over the duty of civility and over rival interpretations of public reason. For an excellent discussion of the reasonable in Rawls (though one inconsistent with some of my arguments in chapter 3), see Wenar, “Political Liberalism: An Internal Critique,”

286 See Rawls, Justice as Fairness: A Restatement, 37: “But given the fact of reasonable pluralism, we try to design our defense of it so as to gain the allegiance of reasonable people and to win wide support.”
incapable of accommodating, let alone favoring, the integrity of persons. All of these serious flaws point to the need for a radical alternative to political liberalism, one that embodies a more plausible conception of respect, a broader grasp of the conditions of political stability, and greater openness to the virtue of integrity as a legitimate aspiration of citizens.

§5.2.1 Respect for Persons

Rawls’s conception of respect, reflected in the so-called “duty of civility” that is supposed to regulate political discourse, brings center-stage the moral requirement to justify coercive interventions in others’ lives—that is to say, more or less deliberate limitations of their autonomy—in terms they could reasonably accept (i.e. accept by responsibly employing their reason), without renouncing their reasonable comprehensive doctrines or converting to another comprehensive doctrine.\(^{287}\) There are three flaws in this distinctively contractualist understanding of respect\(^{288}\): first, if we attend to the actual conditions of political and moral discourse and the normative grounds of coercion, this notion of respect embodies an interpretation of autonomy that is far too stringent to be plausible. Secondly, Rawls’s ideal of citizenship is likely to be politically divisive given its deeply controversial philosophical presuppositions. Thirdly and finally, the contractualist understanding of respect, instantiated chiefly in the principle of reciprocity,

---

\(^{287}\) I explain the distinction between political and comprehensive doctrines in §2.2.4 above. Cf. PL, 13.

\(^{288}\) By “contractualist,” I mean that it participates in the tradition of moral philosophy of Rousseau, Locke, Kant, and Rawls that views political and moral norms as legitimated primarily by the reasonable and rational consent of persons, and conceive of political order as grounded primarily in an actual or hypothetical social contract. I do not mean to suggest that all contractualist theories are committed to such a conception of respect, though as I will argue later, many twentieth century contractualist thinkers, including Habermas, seem to conceive of respect almost exclusively in terms of deference to autonomy. Cf. my discussion of contractualism in §5.3 below.
is overly narrow and impoverished, since it neglects or downplays other crucial dimensions of the concept of respect.

The Implausibility of Rawlsian Respect

Rawls’s understanding of respect, reflected in the duty of civility, is extremely stringent: it requires each citizen to offer others reasons for his or her favored policies and laws that all reasonable citizens\textsuperscript{289}, notwithstanding their conflicting comprehensive doctrines, could conceivably adopt. But in the rough-and-tumble of political debate, whether about detailed policies or about matters of fundamental law and justice, public discourse is littered with appeals to important philosophical assumptions that many reasonable citizens would vehemently dispute. It is enough to think of debates about slavery from the abolitionist era, debates about equality in the American civil rights movement, and current debates about same-sex marriage, to recognize that it is next to impossible to imagine a democratic public sphere cleansed of reasonably contestable assumptions. Or consider the political practice of coalition-building, where one appeals separately to discrete sections of the population, showing each how one’s favored position is supported by their distinctive interests and worldviews.\textsuperscript{290} In itself, this is not a decisive argument against Rawls’s understanding of respect, but it does bring into question its socio-political viability.

\textsuperscript{289} By “reasonable,” Rawls means citizens who are morally and epistemically responsible, i.e. have a sense of justice and employ their reason conscientiously, and (more controversially) accept the duty to offer public or political reasons for their policy proposals, at least when they touch on matters of basic justice or constitutional essentials (see Rawls’s discussion of “the reasonable and the rational” in PL, 48-54). For the purposes of the argument of this section, it makes no difference whether or not we take reasonableness to logically imply adherence to the duty of civility.

\textsuperscript{290} Thanks to the late Phil Quinn for pointing this out during a graduate seminar he led at Notre Dame on the role of religion in public life.
More importantly, it is difficult to see why, under conditions of free and fair inquiry, offering an argument that presupposes deeply controversial and reasonably contestable assumptions is a mark of disrespect for one’s interlocutor. We could readily grant that making little or no effort to render my argument for a law accessible to the other person is a failure to respect him. But this is a far cry from insisting that my argument rely on assumptions that retain their plausibility notwithstanding differences that hold between my comprehensive doctrine and that of my interlocutor. For example, supposing we are arguing over the propriety of a constitutional amendment to enshrine the right of homosexual couples to be married in the eyes of the law. While it may (arguably) be politically imprudent for a citizen to argue for the right to same-sex marriage based on a romanticist view of self-expression and self-fulfillment, given that such a view is not part of the mainstream of our political culture, it is hardly expressive of a lack of respect for her fellow citizens. Similarly, while it may (arguably) be politically imprudent for a pro-life advocate to argue against abortion based on the premise that human life is a gift from our Creator, it is hard to see how arguing from such a controversial premise necessarily shows disrespect for one’s interlocutor. Of course, whenever we profoundly disagree, for whatever reason, with the premises behind an argument, we may be tempted to accuse our adversary of disrespect, but it is crucial that we distinguish between substantive disagreement with our adversary’s argument on the one hand, and disagreement with his style and method of argument on the other.
Of course, we should endeavor, if at all possible, to advance arguments that are rationally accessible to other persons differently situated. But sometimes arguments we consider rationally accessible will be incompatible with others’ deeply held moral and religious convictions. That is a regrettable fact of life, but not a reason for withdrawing the arguments from the political sphere.

The Political Divisiveness of Rawlsian Respect

Rawls’s ideal of citizenship, specifically the “duty of civility,” involves a concept of respect that makes deeply controversial philosophical assumptions, and is thus likely to be too politically divisive to constitute a public guide for civic life. As I have already shown (§3.2.3), the duty of civility presupposes controversial assumptions that are distinctive of a neo-Kantian or contractualist moral epistemology and a neo-Kantian or contractualist moral theory. Let us start with Rawls’s moral epistemology: as I argued at some length in chapter 3 (§3.2.3), one of the premises driving Rawls’s account of moral divergence is a kind of moral egalitarianism that is distinctive of neo-Kantian morality, and would be disputed by many neo-Aristotelians who emphasize the connection between the possession of virtue and the possession of moral knowledge or wisdom.

As I also argued in chapter 3 (§3.2.3), in order to get from a presumption of good faith and equal competency in moral disagreement to the primacy of political over

---

291 Christopher Eberle, for example, while rejecting the strictures of the duty of civility, does believe that respect requires citizens to make a good faith effort to render their arguments accessible to others. See Eberle, Religious Convictions in Liberal Politics.

292 See §5.3.1 below for further explanation of contractualism.
comprehensive justifications, Rawls needs to rely, at least implicitly, upon a reasonably contestable contractualist political morality. Political liberalism, in other words, places a premium on consent to law that is in keeping with a vision of society as a contract among equals. But the primacy of consent in political justification can be readily disputed by citizens from a variety of religious and philosophical persuasions that cannot be ruled out as “unreasonable” except on arbitrary or question-begging philosophical grounds. Our consideration of the epistemological and moral presuppositions of the duty of civility forces us to conclude, given the duty’s pivotal role in Rawls’s ideal of citizenship, that his ideal is more likely to be a source of political division than harmony.

The Narrowness of Rawlsian Respect

In addition to embodying an implausibly stringent understanding of autonomy and being politically divisive, Rawls’s conception of respect has a third flaw that should not be overlooked: namely, that it is severely emaciated insofar as it neglects or downplays many important dimensions of respect. There are many ways of showing respect and disrespect for other persons, and the contractualist imagination, focused as it is on problems of justification and consent, tends not to give other dimensions of respect their full credit. For example, respect in political discourse involves not just offering one’s fellow citizens plausible arguments for one’s favored policies, but doing so honestly, and with courtesy, listening to one’s interlocutor empathetically, having sufficient humility to acknowledge one’s own errors and the merits of one’s political adversary, showing sensitivity to the concerns and interests of the other person, remaining true to one’s word, and giving one’s interlocutor, unless there is strong evidence to the contrary, the benefit
of the doubt. As soon as we pay attention to the sorts of actions and omissions that justify the claim that Tom respects Ellen, we find that respect in political discourse (not to mention other aspects of civic life) is much richer and has many more dimensions than Rawlsian contractualism would suggest. This point will become clearer as I discuss in more detail a richer account of respect rooted in the virtues (in section §6.3.1).

§5.2.2 The Problem of Stability

Rawls clearly aims to uncover the conditions that make for a society in which personal autonomy is respected; but he also aims to uncover the social and political conditions that conduce to political and social stability.\(^{293}\) One of these conditions, according to Rawls, is some public rationale for our fundamental laws that is not purely pragmatic or based on shifting alliances of interests. In order to secure that public rationale, Rawls suggests that citizens observe the “duty of civility,” which imposes a moral duty upon them, when debating constitutional essentials and matters of basic justice, to offer reasons that other (reasonable) citizens could reasonably accept, in spite of their differing reasonable comprehensive doctrines.

The duty of civility may, at least in the short to medium term, foster the appearance of political stability, since it compels citizens to use concepts, evidence, and methods of argumentation that are conventionally deemed “public” or politically accessible. Insofar as the dominant conventions of society concerning what is public and what is nonpublic are stable, and citizens adhere to the duty of civility, the language of

\(^{293}\) As Rawls puts it, “an enduring and secure democratic regime, one not divided into contending doctrinal confessions and hostile social classes, must be willingly and freely supported by at least a substantial majority of its politically active citizens” (PL, 38).
political discourse will be restrained and consistent. However, this veneer of stability is purchased at the cost of genuine, enduring stability. First, a culture in which the duty of civility has taken root will place considerable social pressure upon dissenters to conform, as failure to conform to the duty will be considered the mark of bad or irresponsible citizenship. Given the number of people who would experience the duty of civility as a strain on their integrity, the pressure to conform would likely leave dissenters in the following dilemma: they can either conform reluctantly, and suppress central aspects of their ethical identity, in which case they are likely to build up resentment against a political process that disadvantages them on account of their comprehensive doctrines; or they can participate in the political process but refuse to conform to the duty of civility, in which case they will be perceived as “bad citizens” and will likely resent the slur on their bona fides, and resent a political process that impugns their motives just because of their ethical standpoint. Finally, they may choose to withdraw from a political process that they perceive to be unfairly tilted against their moral and ethical convictions, in which case they will have a reduced stake in the integrity and outcome of the political process. In short, the duty of civility, insofar as it takes root in our political culture, is likely to give rise to a substantial degree of alienation from the political process among citizens upon whom the duty imposes high moral costs. And alienation from the political process, insofar as it weakens the allegiance of citizens to shared political institutions, is an important source of social and political instability.²⁹⁴

Rawls’s fear of reliance upon comprehensive doctrines in political discourse can be explained, in part, by his failure to grasp the major sources of instability in political life. Political instability is not caused principally by philosophical and religious

²⁹⁴ I make this argument in greater detail in §3.4.
disagreements, but by a lack of virtue among citizens, i.e. a failure to habitually act well both towards each other and with respect to the common good of the political community. As I have already explained, Rawls’s account of civic virtue focuses almost exclusively on certain cooperative virtues, above all the master virtue of “reasonableness.” Yet this minimalist scheme of virtue (and by implication, vice) is hardly adequate to identify the major sources of political instability, and the corresponding sources of stability, peace, and good will among citizens. Even a commonsensical survey of political history or of current political events uncovers a broad range of virtues and vices that determine (or partially determine) the quality and outcome of interactions among citizens, many of which simply cannot be captured under Rawls’s “thin” virtue of reasonableness.

For example, in her book *Talking to Strangers* (2004), Danielle Allen talks at great length about the problem of race relations in the United States, and shows how complex the motivations of citizens can be, grounded as they are not so much in commitments to abstract norms of behavior as in particular human relationships of trust or distrust, admiration or suspicion, and in particular character traits, whether virtues such as courage and patience, or vices such as cowardliness, pusillanimity, or arrogance. She gives the example of that historic confrontation between State troopers and white parents on the one hand, and a black student, Elizabeth Eckford on the other, in Little Rock, Arkansas, in 1957. The tensions between blacks and whites on that occasion did not arise merely from the unwillingness of whites to offer fair terms of cooperation, or their incivility towards blacks, but from a history of oppression marked by acts of

---

295 I might never have discovered this book, which I found extremely insightful for my project, had Jeffrey Stout not recommended it to me.
arrogance, dishonesty, cowardliness, pettiness, contempt, and profound distrust. Rebuilding relations of trust and respect between blacks and whites would be not so much a question of observing certain norms such as the “duty of civility” as building up a friendship by establishing one’s *bona fides* through the consistent display of virtues such as humility, honesty, trust, courage, and magnanimity. A richer account of civic virtue than political liberalism can provide is required in order to explain the conditions under which trust grows or wanes between citizens. These points will hopefully become clearer in chapter 5 where I outline a virtue ethical ideal of citizenship and respect and show how it can help us tackle the problem of political stability better than Rawls’s ideal.

§5.2.3 The Problem of Integrity

As I explained in chapter 2, Rawls’s ideal of citizenship imposes a duty upon all citizens to rely upon public or political values while debating matters of basic justice. But upon examination, we found that his interpretation of the publicity requirement, which required an appeal by citizens to reasons “any reasonable person could accept,” would impose a severe burden upon many responsible and cooperative citizens otherwise committed to liberal values and institutions, who would feel forced to choose between remaining faithful to their deepest ethical commitments and thus living with integrity; and adhering to the political liberal ideal of citizenship, and thus living up to their (putative) public duties. Yet the imposition of such a burden, as I argued at some length in chapter 3, is neither politically practical nor philosophically warranted. And as I will argue shortly (chapter 6), there are better ways of grounding stable civic interactions that do not rely upon such a philosophically partisan and exclusionary ideal.
§5.3 The Limits of Contemporary Contractualism

§5.3.1 John Rawls: Political Liberalism

As I have mentioned in passing a number of times, Rawls’s project is contractualist at its core, that is to say, it is inspired by a vision of society as a contract among equals, and reduces political morality to that which can be agreed upon by suitably situated or “reasonable” agents, whether actual or hypothetical. As part of my diagnosis of the failure of political liberalism to meet the challenge of modern political order, I want to suggest—though for now it must remain little more than a plausible suggestion—that the limitations of political liberalism are attributable, in no small part, to the contractualist framework which structures its conceptual world.

What exactly do I have in mind here? I do not have in mind the entire tradition of social contract theory going back to Hobbes and Locke. Instead, I have in mind the more recent form of social contract theory that we have seen emerging in the twentieth century, with works such as Rawls’s Theory of Justice (1971), Nozick’s Anarchy, State, and Utopia (1974), Scanlon’s What We Owe to Each Other (1998) and Habermas’s Theory of Communicative Action (1984). This form of social contract theory focuses on the rational justification of political norms to rational subjects as the essence of constituting modernity, see Taylor, Modern Social Imaginaries.

Indeed, early modern liberals such as Hobbes and Locke, while placing consent at the center of their theories of political order, were much more attentive than most of their successors to the central role of virtues and mores in sustaining a liberal regime. For a careful survey of the role of virtue in the liberal tradition, see Berkowitz, Virtue and the Making of Modern Liberalism.

political morality, and relegates broader issues of character, virtue, and ethical flourishing to the realm of “ethics” or to the “background culture,” not considering these issues the concern of the political philosopher as such. Political morality or normative theory is concerned, on this contractualist model of politics, almost exclusively with winning the rational (or reasonable) consent of citizens to the political institutions that govern their lives.

Now, if you consider the relatively narrow approach to political morality and legitimacy embodied in the notion of a social contract, it makes sense that somebody guided by that tradition would neglect a range of conceptual possibilities that cannot be described in contractualist terms. I want to focus on two such possibilities here: first, the possibility that political respect is a complex attitude informed by many factors only one of which is deference to autonomy; and second, the possibility that stability depends at least as much on the virtues and vices reflected in citizens’ behavior as on the availability of compelling justifications for policies and laws. If you adopt a contractualist approach to political order, respect tends to be reduced to respect for rights or respect for law, or respect for the autonomy of others, since a contract, of its very nature, is founded on consent and respect for the freedom to enter and exit contracts voluntarily. Yet as an analogy for political society, the notion of a contract is extremely limited, since contract is a creature primarily of the legal order, whereas political society is a creature of the

---

299 Think of Habermas’s distinction between the “ethical” and the “juridical” spheres, and Rawls’s exhortation to apply the “principle of toleration to philosophy itself” (Rawls, "Justice as Fairness: Political Not Metaphysical," 231).

300 The “background culture” is distinguished from the public or political culture by Rawls.
social and cultural order and the mores of the people, which extend far beyond respect for autonomy or consent, as Tocqueville and Montesquieu could tell us.\(^{301}\)

A contractualist thinker will tend to think of political stability as a function of rational consent: just as voluntary contracts do not bind unless people are rationally convinced to enter into them, so, on the contractualist account, the grand “contract” of political society cannot bind citizens to itself unless it is presented in terms that are rationally cogent and persuasive to them. Yet this neglects a range of factors that inform the social and political cohesion of a society, including the virtues and vices manifested in citizens’ mutual treatment and conversations, and the capacity of a range of intermediate associations to foster the virtues that conduce to political stability. To clarify, I am claiming neither that rational argumentation is irrelevant to stability, nor that social contract thinkers are altogether blind to the virtues and vices that inform the stability of a regime. I do want to claim, however, that social contract thinkers tend to emphasize the role of procedures, norms, and institutions, to the neglect of the role of civic virtue in sustaining a liberal regime; or where they do treat of civic virtue, it tends to be fairly narrowly conceived as something oriented towards rational argumentation about law and policy rather than towards the good of other persons or the overall health of the polity. These tendencies seem perfectly understandable to me within a tradition that thinks of society as analogous to a legal contract, rather than, say, a family or a village.

---

that evolves over time based on the peculiar combination of nature and nurture that constitutes the human condition.\(^{302}\)

Faced with the inadequacy of political liberalism across the axes of respect, stability, and integrity, we must seek elsewhere for an ideal of citizenship that satisfactorily addresses the problem of political order in an environment of deep-seated religious and moral pluralism. In order to bring into focus the distinctiveness and broader significance of the aretaic alternative to be advanced in the next chapter, it is worth examining two other ideals of citizenship that are both critical of political liberalism and represent important trends within contemporary political theory: deliberative democracy and agonistic pluralism. I want to argue that notwithstanding the distinctiveness of these ideals, they too fail to convincingly address the questions of respect, stability, and integrity in contemporary political life. I will further argue that at least some of their limitations stem from the contractualist framework they share with Rawls.

For ease of exposition, I will choose two authors who are readily recognizable exponents of both views, Irish Young and Chantal Mouffe. My choice of these authors is not meant to reflect on the quality or significance of other works but instead reflects my judgment that Young and Mouffe provide representative and relatively strong samples of two of the currently dominant strands of political theory. Furthermore, I would contend that the limitations I identify in Iris Young’s deliberative democracy and Chantal Mouffe’s agonistic pluralism are representative of limitations within the currents of political theory that each participates in rather than idiosyncratic difficulties with their particular articulations.

\(^{302}\) See, for example, Aristotle’s description of the evolution of the *polis* from the household and village in *Pol. 1*. 
§5.3.2 Iris Young: Deliberative Democracy

Iris Young shares my misgivings about Rawls’s norm of “reasonableness,” in particular its exaltation of civility and consensus at the expense of democratic dissent, and its overly narrow and rationalistic construal of civic discourse or communication as the formal exchange of reasons and arguments. In her book *Inclusion and Democracy* (2000), Young advances what she takes to be a more open, inclusive, and genuinely pluralistic vision of civic life, one that rejects the exclusionary implications of Rawlsian reasonableness and seeks to accommodate the full array of perspectives and styles of communication at large in democratic polities. Young understands herself to be advancing a variant of “deliberative democracy,” a relatively recent strand of political theory that emphasizes the deliberative and transformative, rather than the procedural and a prioristic, aspects of democracy.  

Somewhat like Rawls and other recent contractualist thinkers, Young tends to focus on institutional and normative questions and does not spend a lot of time developing an ideal of civic character per se. However, there is an ideal of citizenship latent within her account, and I want to argue that though it escapes

---


304 The priority of institutional and normative issues over issues of character can be observed in thinkers as diverse as John Rawls, Jürgen Habermas, Iris Young, Chantal Mouffe, Stephen Macedo, and Philip Pettit, all of whom I take to be contractualist in the sense I have stipulated.
some of the flaws of political liberalism, it is susceptible to the same basic critique I have advanced against Rawls.

Young’s model of civic interaction accepts the basic features of most accounts of deliberative democracy, which she outlines in chapter 1 of *Democracy and Inclusion: inclusion and political equality; reasonableness; and publicity*. According to the norm of inclusion and political equality, nominal inclusion in the process is not sufficient: all those included must be included “on equal terms,” and be afforded “an equal right and effective opportunity to express their interests and concerns.” The norm of reasonableness entails a “willingness to listen to [and be corrected by] others.” Reasonable people aim at reaching agreement even if they fail to attain this goal. Finally, reasonable people “must have an open mind,” meaning that “[t]hey cannot come to the discussion of a collective problem with commitments that bind them to the authority of prior norms or unquestionable beliefs.” *Publicity* entails a commitment to express oneself in a way that “aims in its form and content to be understandable and acceptable,” and accountable to the pluralistic or diverse audience that constitutes the modern polity.\(^\text{305}\)

Although Young “endorses the basic outlines” of the deliberative model as formulated here, she does see the need to qualify it in several respects: (i) first, *pace* Habermas, particular or “partial” interests must be *engaged* rather than set aside in political discourse; (ii) second, no “common good” but rather an arrangement of interests, is the aim of politics; and (iii) third, norms of political communication must acknowledge the

\(^{305}\) See Young, *Inclusion and Democracy*, 23-25.
legitimate role of less rationalistic forms of communication such as protest and less “calm,” more emotionally driven forms of speech.  

Young’s ideal of citizenship is in some respects a significant improvement on Rawls’s ideal: for example, Young’s account draws attention to the role of less discursive or “rationalistic” forms of communication such as emotionally charged rhetoric and disruptive protests, in advancing democratic goals. Also, she is more sensitive than Rawls appears to be to the exclusionary tendencies of even ostensibly “open” political processes, e.g. the way the language and procedures of a public forum can wittingly or unwittingly exclude women, minority cultures, etc. Thirdly, she offers a promising version of the publicity requirement that replaces Rawls’s unduly stringent reciprocity test (public reasons conceived as reasons any reasonable person could accept) with the morally substantive yet realistic requirement that the expression of one’s arguments “aims in its form and content to be understandable and acceptable” (25, my emphasis).

Notwithstanding these improvements, Young’s theory of deliberative democracy falls prey to the same basic limitations I pointed out in Rawls’s account. First of all, like Rawls’s notion of respect, Young’s is philosophically implausible and too “thin” to provide substantial guidance to citizens in their mutual interactions. Secondly, in spite of her ideal of inclusiveness, in reality her ideal of reasonableness is a heavy and unwarranted burden on the integrity of many citizens. Finally, the politically divisive nature of her ideal and her inattention to the broader range of civic virtues disables her account from promoting or sufficiently illuminating the conditions of political stability.

---

306 This is a highly abridged version of Young’s account, and though it leaves out many details, I believe it is accurate and faithfully conveys the gist of her argument.
Insofar as the limitations of Young’s account resemble those of Rawls, this can be explained, at least in part, by the spirit of contemporary contractualism that shapes it.

Consider first Young’s notion of respect: I would begin by saying that I have no quarrel in principle with the notion of political equality and inclusiveness. However, I do find Young’s interpretation of this ideal problematic. On her account, “reasonableness,” which I take to be an aspect of respect (though not much hangs on this particular assumption), implies that citizens “cannot come to the discussion of a collective problem with commitments that bind them to the authority of prior norms or unquestionable beliefs” (24). Yet it is difficult to know how any meaningful normative debate could proceed without parties to the discussion feeling “bound…to the authority of prior norms or unquestionable beliefs.” Indeed, it seems that bedrock moral commitments form the necessary presupposition of any moral conversation, or at least any constructive moral conversation in a democratic society. Some beliefs are simply beyond question—they cannot be debated as if they were subject to revision. They include the belief that human beings are owed respect or have inherent worth; the belief that survival is preferable to mass suicide barring extraordinary circumstances; and the belief that torturing children for fun is morally heinous. I simply cannot see how “reasonableness” would require citizens to unbind themselves from the “authority” of such “prior norms.” In short, at least this aspect of Young’s concept of respect does not seem like a promising basis for a shared ideal of citizenship in democratic societies, either on pragmatic or moral grounds.

Aside from the implausibility of Young’s norm of “reasonableness,” her general understanding of respect, though it has laudable features such as openness to diverse perspectives and modes of communication, is too thin to provide substantive guidance to
citizens in their daily interactions. Like Rawls’s understanding of respect, Young’s is quite formal, involving notions such as fairness, inclusiveness, and reasonableness, but it overlooks a broad range of dispositions that may express an attitude of respect in different social contexts, e.g. honesty, truthfulness, humility, dependability, compassion, justice, and magnanimity.

As a consequence of the implausibility and divisiveness of Young’s notion of respect, it may well impose philosophically unwarranted burdens on the integrity of many citizens, at least insofar as it becomes established as a norm of public discourse. After all, it seems safe to assume that most citizens do in fact adhere to “unquestioned” “prior” normative beliefs, whether they are aware of their adherence or not. Those who are aware of such adherence are likely to feel alienated from a political culture that embodies the requirement to set such commitments aside upon entering the political process.

This bring us to the final problem with Young’s ideal of citizenship: its incapacity to satisfactorily address the challenge of stabilizing interactions within a pluralistic society, or at least shedding light on the conditions necessary for stability. The first and most obvious point to note in this regard is that, as I have just argued, Young’s norm of reasonableness is roughly as politically divisive (perhaps even more so) as Rawls’s duty of civility, and thus cannot serve as a common moral reference point for diverse citizens. The second and more subtle point to note is that Young’s account of citizenship, because of its bias towards institutional and normative issues, tends to neglect the central role of virtues or character dispositions in stabilizing civic interactions. There are many virtues,

---

307 I am not suggesting that a concept of respect can serve as a rule to be applied deductively to practical problems. I think of it more as a regulative ideal that shapes and directs the spirit of civic interaction.
as I pointed out in my response to political liberalism, that contribute to political stability, and Young, like Rawls, is insufficiently attentive to them, e.g. courage, conscientiousness, dependability, humility, magnanimity, generosity, etc. This point should become clearer in light of my positive account of civic virtue in chapter 6.

Finally, I have suggested that some of the limitations inherent in Young’s account, in particular those that resemble the limitations of political liberalism, are attributable in part to its contractualist spirit. By this, I mean that Young’s account, like Rawls’s, tends to think of society primarily as a contract among equals, rather than an alliance of “nature and nurture,” and conceives of political morality and legitimacy primarily as an agreement among rational actors on a set of rules that constitute the social contract. This contractualist spirit leads Young, like Rawls, to attend to normative and institutional questions to the neglect of questions of character, and to focus on a fairly narrow range of virtues associated with political bargaining and argumentation, such as reasonableness, publicity, and inclusion. Finally, the notion that morality is constituted rather than discovered by rational agreement, which seems central to the most dominant forms of twentieth century contractualism, leads Young to implicitly exclude any constitutive, unquestionable, a priori, or argument-transcending claims from political discourse.

---

308 I have in mind works such as John Rawls, A Theory of Justice (Oxford: Oxford University Press, 1971); Habermas, The Theory of Communicative Action; and Scanlon, What We Owe to Each Other.

309 I hasten to add that though most of us do ground our discussions in some a priori or unquestionable assumptions, the fact that a particular person treats claim X as unquestionable or beyond compromise does not automatically validate claim X or disbar other discussants from bringing it into question.
§5.3.3 Chantal Mouffe: Agonistic Pluralism

Chantal Mouffe, like Iris Young, is suspicious of consensus-driven theories of political discourse, and wants to widen the field of discourse to include diverse voices and styles of argumentation. She rejects as illusory the claims to normative or philosophical impartiality of both John Rawls’s political liberalism and Jürgen Habermas’s discourse ethics. And she points out that conflict is an ineliminable and central dimension of politics, which both political liberalism and deliberative democracy tend to overlook or suppress in the name of universalizing norms of rationality. I have no argument in principle with any of these claims; indeed, I consider them an important correction to the simplifying and reductive tendencies of both political liberalism and deliberative democracy. However, I do not believe Mouffe fares much better than either Rawls or Young in offering a defensible notion of respect, a political morality sensitive to ethical integrity, or a plausible basis for tackling the problem of political stability. And surprising as it may seem at first, Mouffe, like Rawls and Young, is driven to her conclusions in part by her immersion in the contractualist mentality.

Mouffe’s favored approach to politics is what she calls “agonistic pluralism” and it is laid out in her books *Hegemony and Socialist Strategy* (co-authored with Ernest Laclau) and *On the Political*, and summarized in an article entitled “Deliberative Democracy or Agonistic Pluralism?”310 This approach rejects both the neutralist proceduralism of individualist liberalism, where interests are formed prior to discourse, and the moral rationalism of deliberative democracy, which assumes that we can

transcend conflict and partisanship by adhering to universal and shared moral principles. Mouffe argues that both the “moral” and “economic” approaches to political discourse neglect the dimension of the “political” which she defines as “the dimension of antagonism that is inherent in all human society.”\textsuperscript{311} Agonistic democracy does not aim to suppress or eliminate this antagonism, but to “domesticate” it by conceiving our adversary not simply as an enemy to be eliminated, but as “a legitimate enemy, an enemy with whom we have in common a shared adhesion to the ethico-political principles of democracy.”\textsuperscript{312}

The problem with Mouffe’s approach is that for all its celebration of diversity and pluralism, it never makes clear what the normative grounding for civic interaction is, or why diversity is a good to be valued or respected. On the one hand, Mouffe never specifies clearly what the “ethico-political” principles of liberal democracy are. On the other hand, even if she did, there would be no reason to think of them as anything more than the arbitrary constructions of political power, since on her view, “social objectivity is constituted through acts of power” and “any social objectivity is ultimately political.”\textsuperscript{313} Though Mouffe is right to highlight the ineliminability of conflict in politics, she ends up reducing political discourse to a power play with no objective moral grounding whatsoever. Political legitimacy is no longer conceived as a morally intelligible and arguable claim, but as a claim based on “purely pragmatic grounds.”\textsuperscript{314}

\textsuperscript{311} Mouffe, “Deliberative Democracy or Agonistic Pluralism?,” 754.

\textsuperscript{312} Ibid., 755.

\textsuperscript{313} Ibid., 752 (emphasis added).

\textsuperscript{314} Ibid., 753.
Ethics and morality are not only obscured by, but are collapsed into pragmatism and power politics.⁴¹⁵

It should be clear by now that Mouffe’s account, because of its elision of power and morality, cannot provide any clear or cogent grounds for distinguishing between respectful and disrespectful political engagement. In fact, it is not clear that there is room in Mouffe’s account for any recognizable concept of respect—that is, a concept of respect that draws non-arbitrary normative limitations to the mutual relations between human beings. Having built in a formidable obstacle to any non-pragmatic concept of respect, it is not surprising that Mouffe’s view of citizenship does not say much about what respecting a fellow citizen entails, beyond alluding from time to time to the “ethico-political principles of democracy.” Thus, her account is not only ill-suited in principle to the notion of respect: it also sheds little substantive light on what respect involves in civic interactions.

With regard to shedding light on the problem of stability, Mouffe’s account does not fare much better. While she does well to highlight the agonistic or conflictual dimension of politics and to draw our attention to the danger inherent in a consensual vision of political order, there is ultimately nothing in Mouffe’s account to stabilize the plurality of perspectives, beyond “acts of power” which “[constitute] social objectivity.”³¹⁶ Thus, to the extent that stability is obtained, it is apparently purchased by the imposition of power. While this may be how stability is secured in many situations, it is unsatisfactory from a broadly liberal perspective, since it effectively separates political

---

³¹⁵ It is telling that Mouffe herself uses this terminology, when she points to the “convergence” or “mutual collapse…between objectivity and power” (752-753).

³¹⁶ See Mouffe, “Deliberative Democracy or Agonistic Pluralism?,” 752.
stability from questions of legitimacy and human flourishing. Stripped of its ethical underpinnings, stability is no more than the arbitrary imposition of power. This sombre view of political stability, perhaps even more than political liberalism and deliberative democracy, seems to overlook a range of virtues that contribute to the stability of discourse and social interaction—virtues to be explored in greater detail in chapter 5.

On the issue of integrity, agonistic pluralism, while it opens up a space for self-expression in the political arena, neither advances a coherent vision of human flourishing nor provides a vision of politics compatible with a genuine ethical discourse among citizens about their common life together. Agonistic pluralism simply does not offer a sufficient bridge between politics and ethical agency or human flourishing to make personal integrity a realistic aspiration for citizens. On Mouffe’s account, it is not clear how a citizen’s personal ethical commitments could, or should, inform her political interactions. Indeed, one implication of Mouffe’s agonistic model of politics would appear to be that we are driven back to a rigid separation between the private realm, where we construct ethical lives as best we can, and the public realm, where we participate in a strategic struggle of interests with our political adversaries against the backdrop of norms that are constituted by the raw exercise of power.

Finally, I want to suggest that Mouffe’s agonistic conception of politics is shaped to a large extent by its immersion in contemporary contractualism. At first sight, this claim may seem obviously mistaken, given the fact that contractualism emphasizes moral rules as the outcome of universal agreement, while agonistic pluralism, on the contrary, rejects universal agreement as a convenient and dangerous illusion. However, upon closer examination, Mouffe’s view of political morality turns out to be the mirror image of the
contractualist view, and in that sense remains captive to the contractualist imagination. Where contractualism posits morality as the outcome of a universal agreement, Mouffe, arguing within the same basic model, rejects the agreement as an illusion, and moves quickly to dismiss the validity of moral norms as anything more than socially constructed. Where there is no rational agreement, there is no objective morality. That is a premise Mouffe inherits from contractualism—no Aristotelian or Thomist, for example, would accept such a premise. Similarly, just as contractualists tend to associate political stability almost exclusively with philosophical agreement on the ultimate ground rules of politics, Mouffe, in recognizing the instability and often absence of such agreement in contemporary politics, quickly infers that politics is little more than a thinly disguised struggle for power. Where there is little rational agreement on the terms of the contract, we are thrown back on a quasi-Hobbesian passion for political power and self-assertion, limited only by a circumscribed set of “ethico-political principles” upon which many of us can agree—the social contract reintroduced by the back door, so to speak. The essential point is that Mouffe’s account of politics is still deeply influenced by the contractualist paradigm, which posits universal (or quasi-universal) agreement on principles rather than the possession of virtues, as the main source of political stability, moral truth, and mutual respect.

The basic goal of this chapter has been to deepen and broaden my diagnosis of the failure of political liberalism to convincingly address the problem of moral and religious pluralism. I have done this by setting out in concise form three basic criteria that any successful account of citizenship must meet from a broadly liberal perspective; arguing that political liberalism fails to meet these requirements; and showing how this failure
can be traced, in no small part, to the influence of contemporary moral contractualism. To support this claim further, and widen the significance of my critique of political liberalism, I have examined two significant rivals to political liberalism, deliberative democracy and agonistic pluralism, and shown that they exhibit related if not identical limitations that can be explained in part by the pervasive influence of twentieth century contractualism. In the next chapter, I will set out some of the basic features of an alternative account, which I will call the *aretaic* ideal of citizenship, and argue that this account rivals and surpasses the three contractualist theories we have considered in this chapter along the axes of respect, stability, and integrity.
CHAPTER 6: TOWARDS A POLITICS OF INTEGRITY AND

VIRTUE: THE ARETAIC IDEAL OF CITIZENSHIP

We are now approaching the end of this investigation of the ethical foundations of modern citizenship. Recall that the goal of this investigation was to provide some grounds for the hope that one could conscientiously pursue the good life according to one’s best lights, while remaining a citizen in good standing in a liberal democratic polity. We examined this hope from two principal perspectives: first, looking at modern citizenship from the perspective of ethical integrity, to see whether it could be embraced without betraying one’s deepest ethical commitments; and second, looking at integrity from the perspective of liberal democracy, to see whether it could be accommodated and supported without undermining or even jeopardizing liberal institutions, values, and practices. We found that the most dominant theoretical accounts of modern citizenship are indeed in sharp opposition to personal integrity, and that they lack the resources for any substantial rapprochement between political order and the quest for the good life. Yet we also found, based on a cursory examination of the practice of modern citizenship, that a mature account of integrity seemed in many respects compatible with modern political order. Based on that preliminary finding, and on the limitations of current theories of citizenship, this final chapter plumbs the resources of classical virtue theory to test the hypothesis that an aretaic ideal of citizenship might hold out the hope for some level of
reconciliation between the demands of political order in a liberal regime on the one hand, and the demands of ethical integrity on the other.

(§6.1) I will frame my account of citizenship by providing a rough conceptual framework for understanding the ethical significance of citizenship, or its relation to a person’s pursuit of the good. (§6.2) I will then present the broad contours of a virtue-based ideal of citizenship friendly to integrity yet tailored to modern socio-political conditions, and elaborate this aretaic vision of citizenship in the crucial realm of political discourse. (§6.3) Thirdly, I argue for the superiority of the aretaic317 ideal of citizenship over the political liberal, agonistic, and deliberative models, from a broadly liberal perspective318, along three axes which correspond to my criteria (stated at the outset of chapter 5) for a successful account: (a) respect for persons (§6.3.1); (b) political stability (§6.3.2); and (c) the demands of integrity (§6.3.3). (§6.4) Finally, I canvass seven possible objections against the aretaic ideal of citizenship and show that they rest on misunderstandings or distortions of various aspects of the aretaic model.

§6.1 Personhood, Citizenship, and the Good Life

The fundamental question that goes to the heart of political philosophy, yet is rarely raised overtly by contemporary political philosophers, is: “Is it possible to be a

---

317 I use the terms “virtue-ethical” and “aretaic” interchangeably. “Aretaic” is used much less frequently by philosophers than “virtue-ethical” but it is stylistically more elegant and conveniently serves to remind us of the Greek origin of the concept of virtue.

318 As I point out later, by broadly liberal, I do not mean Rawlsian liberal, neo-Kantian, or indeed Lockean, but rather, a philosophical outlook that accepts in principle the legitimacy and goodness of the fundamental institutions and practices characteristic of liberal regimes, such as private property and the free market economy, personal liberty within the limits of the common good, equality before the law, freedom of religion, freedom of conscience, freedom of association, rule of law, democratic procedures, and the right to due process.
good citizen and a good person simultaneously? How does my citizenship bear on my efforts to lead a good life, and vice versa?” Indeed, this more than any other question is the motor of this entire dissertation. Liberal thinkers such as Rawls systematically avoid it, or only treat it in a very formal manner, since they are precluded by their commitment to ethical minimalism from dabbling in questions that would stray too much into the arena of personal, “non-political” ethical values or goods. Yet by adopting this strategy of avoidance, liberalism robs itself of the resources necessary to offer a convincing vindication of its institutions and practices to those who doubt the compatibility of the liberal democratic ethos with a truly well-lived human life. Furthermore, the strategy of avoidance has led many liberals to overlook the ethical resources a richer account of human virtue can offer to political philosophy, both concerning the virtues that contribute to political stability, and the virtues that constitute an attitude of respect between citizens.

I have already attempted to paint a rough picture of an integrated human life, or a life that has achieved the fullness of integrity (chapter 4). In addition, I have responded briefly to a number of objections to the virtue of integrity, in part in order to show that it can be construed in a way that is prima facie compatible with responsible citizenship and the modern proliferation of roles. However, my defense of integrity is not complete until I have fleshed out an aretaic ideal of citizenship in more detail, and indicated how such an ideal can be both responsive to liberal democratic values, and compatible with

---

319 The notion of a “minimal” moral foundation for politics is suggested by Larmore in his essay, Larmore, “The Moral Basis of Political Liberalism.”

320 Liberal thinkers such as Ronald Dworkin and William Galston have voiced the same sort of concern, though their answers to it differ from mine substantially. See, inter alia, William Galston, Liberal Purposes: Goods, Virtues and Diversity in the Liberal State (Cambridge: Cambridge University Press, 1991); Galston, Liberal Pluralism; Dworkin, “Foundations of Liberal Equality”; and Tomasi, Liberalism Beyond Justice.
integrity as I understand it. The starting-point for this general aretaic vision of citizenship is the concept of the person and the connection between human goodness and the virtues. This discussion will serve to frame and illuminate the role of citizen and the notion of civic virtue encompassed by that role. I will argue that properly understood, the role of citizen, while it does generate special reasons for action where that role is in play, is similar to other human roles in that it is only intelligible insofar as it is constituted by the decisions, commitments, and nature of a person qua person. A role may be shaped by the peculiarities of a culture, regime, or practice, but it must find its place in a good human life to be worthy of a person’s allegiance.

Concomitantly, civic virtue—a central aspect of the role of citizen—cannot be understood in isolation from, or independently from, the wider canon of ethical or moral virtues. Although it is legitimate to distinguish the personal ethical standpoint from the public political standpoint, most of the recent literature on civic virtue has tended to adopt the public standpoint on civic virtue to the neglect of the personal.\(^\text{321}\) I would contend that any morally cogent account of citizenship must show that civic virtue is, at least aspirationally if not always in practice, a part of personal virtue simpliciter rather than a morally indifferent, entirely regime-determined category of virtue.

Let us start, then, with the question of citizenship and the good life. Is it possible to be a good person and a good citizen of a liberal democratic regime such as those of Europe and North America in the year 2005? Or must we choose between them? Some,

\(^{321}\) One of the most famous, or infamous classical instances of the priority of the public interest and authority over ethics is of course Hobbes’s *Leviathan*, which considers personal ethics almost exclusively as a function of public order and stability. More recently, authors who emphasize the public standpoint and pay little attention to the personal standpoint in politics include John Rawls, Charles Larmore, Stephen Macedo, Jürgen Habermas, Benjamin Barber, William Galston, and Peter Berkowitz. Of course, I do not mean to suggest that they systematically subordinate the personal to the political as Hobbes did.
such as John Howard Yoder, virtually repudiate the role of citizen as being incompatible with the dictates of the good life, as understood from a particular Christian standpoint.\footnote{See John Howard Yoder, \emph{The Original Revolution} (Scottdale, Pennsylvania: Herald Press, 1972).} Others, such as Stephen Macedo, are more sanguine, affirming a broadly liberal conception of human well-being and finding, not surprisingly, that liberal citizenship promotes the good life so understood.\footnote{See Stephen Macedo, \emph{Liberal Virtues: Citizenship, Virtue, and Community in Liberal Constitutionalism} (Oxford: Clarendon Press, 1990). Along similar lines, see Galston, \emph{Liberal Purposes}.}

In asking whether citizenship in a liberal regime can be squared with a truly \textit{good} human life, one of the first questions to be addressed is of course, what \textit{is} a “good” human life? Although no philosopher worthy of the name would claim to provide a definitive and exhaustive account of human goodness (for then, practical philosophy might become simply a matter of elaborating on one philosopher’s findings!), we cannot rest content with a purely formal definition: some content must be provided, if we are to avoid remaining at the level of vague abstractions. Rather than provide an ambitious or far-reaching account of a well-lived human life, I want to focus on one of the distinguishing features of an admirable human life, and consider it a central necessary condition for a good life, though it is by no means a \textit{sufficient} condition: integrity. As I have already proposed in chapter 4, integrity is a family of virtues or good-making dispositions that an agent must possess in order to (a) have a persistent highest-order desire to do what is (objectively) good; (b) take appropriate steps to determine what is (objectively) good; and (c) order or re-order his moral life (beliefs, evaluations, desires, dispositions, actions) over time in accordance with that determination. Some of the virtues most closely associated with integrity, as I argued in chapter 4, are temperance,
honesty, perseverance, conscientiousness, humility, and courage. Even if we accept MacIntyre’s rather open-ended definition of the good life as “the life spent in seeking for the good life,”\textsuperscript{324} integrity will be one of the defining features of such a life. So, whatever other virtues a good person may have, they must include the virtues of integrity. Consequently, a basic threshold that any ideal of citizenship must meet, if it is to recommend itself to the moral sensibilities of persons who seek to live well, is that it can accommodate the virtues of integrity.

Before cataloguing some of the dispositions or virtues that constitute the aretaic ideal of citizenship, we must say something about four terms that provide the basic conceptual framework for the ideal: (1) the \textit{role} of citizen which encompasses (2) \textit{civic virtue}; and two correlative terms, which are often neglected in the literature on citizenship, namely (3) the concept of \textit{person} which encompasses (4) \textit{personal virtue}. My central claim, that the aretaic ideal is more friendly to personal integrity than political liberalism and other contractualist accounts of citizenship, will be advanced by offering a distinctive interpretation of the relation between \textit{personhood}, \textit{personal virtue}, the \textit{role} of citizenship, and \textit{civic virtue}. Because this approach to citizenship is not typical among theorists of citizenship who often have a strong bias towards institutional and normative questions\textsuperscript{325}, some of my foundational remarks (e.g. on the nature of a person; and on the notion of a role) may appear beside the point. However, it should help the reader if she keeps in mind as I lay the conceptual groundwork for my ideal, that the fundamental

\textsuperscript{324} See MacIntyre, \textit{After Virtue}, 219.

\textsuperscript{325} As I suggested earlier in my critique of twentieth century social contract theory, the contractualist approach to political order, seeing political society through the analogy of \textit{contract}, understandably focuses heavily on the legalistic and institutional aspects of politics rather than the social, cultural, and dispositional or aretaic determinants of political order.
question driving my account is not merely, “what is the role of citizen and what are its requirements?” but, “how, if at all, can the role of citizen within a liberal democratic polity fit into a good human life (or, to use MacIntyre’s formulation, a life spent in pursuit of the good)?”

§6.1.1 The Concept of Person and Personal Virtue

Let us begin with the closely associated concepts of person and personal virtue. The reason I begin with these concepts rather than with the concept of citizenship is that they are both morally and ontologically more fundamental, and are always in the background, irrespective of the political or social context in which we find ourselves. We may from time to time have a doubt about whether, or in what sense, we are citizens—indeed, in extreme cases such as expulsion from one’s native land, one may be “stripped” of one’s citizenship—but our personhood, that is, the fact that we are human beings with personal attributes, is something we retain even in the most abject of circumstances, such as a prison camp, where our citizenship seems tenuous if not irrelevant. That is not to say that we cannot be subject to conditions that impair or even eliminate our distinctively human capacities or activities, e.g. severe brain damage, but in the vast majority of cases, humans do possess, either latently or actually, the capacity for distinctively human activities such as love, critical knowledge, self-awareness, self-conscious willing, and second-order desires (i.e. desires about our desires). For the purposes of this account, I

326 I am not suggesting by this remark that those who do not possess such attributes either latently or actually are non-persons. I share the belief of many philosophers that a minimal requisite of personhood, at least in the visible world as we know it, is membership in the species, homo sapiens. Beyond that, there is no need for me to enter here into debates about the moral status of embryos and infants, the mentally incompetent, or the disabled, important as these questions are. Our primary concern in this dissertation is with those persons who possess some capacity—however primitive or undeveloped—to deliberate and choose, and thus can, at least potentially, exercise the role of citizenship and its associated virtues.
will assume that humans, just insofar as they possess the capacity for those personal activities that mark out humans from other animals, can be considered persons. Thus, persons and humans, though conceptually distinct, may turn out to be coextensive, depending on how wide one’s reading of “capacity for personal activity” is (e.g. does a human at the foetal or embryonic stage have the “capacity” for personal activity?).

As we have already seen in chapter 2, citizenship is a role which is constituted in large part by a special relationship that persons enter into (or find themselves in) with each other. It is, to a significant degree, socially constructed, yet for all that, it is a person who occupies the role, or who contracts the relationships that distinguish that role, and it is other persons to whom we are accountable for our exercise of the role, rather than other roles or other persons viewed solely as role-holders (whatever that would mean). If this were not the case, we could not explain how it is possible to offer normative arguments against role behavior that is “faithful” to the role of citizen within the regime in question, yet seems to demean or mistreat human persons qua persons. For example, if a regime is premised on the annihilation and/or expulsion of the Jewish race, and citizenship is conceived within the regime as perfect collaboration with that goal, one may be expected or even required, from an ethical perspective, to repudiate the role of citizen in order to avoid demeaning other persons (and in the process sacrificing the integrity of one’s own person).

In other words, the role of citizen is the role of a person, and a person is ethically accountable as a person for the way in which he or she exercises his or her social roles.

However, see MacIntyre, Dependent Rational Animals, for a welcome correction of the tendency among political and moral philosophers to exalt human autonomy to the utter neglect of human dependency and the “virtues of acknowledged dependence.”
This is not to suggest that the evaluation of role behavior can be derived exclusively from role-independent ethical standards—after all, roles involve particular relationships and responsibilities that require prudence to be correctly interpreted. Nonetheless, we cannot “hide” behind our roles and pretend that we have “no choice” but to enact them. The very process of justification or rationalization of our behavior is a testament to the fact that we can affirm or negate the requirements of a role if they contradict the goods we are committed to pursuing and instantiating as persons. This is not to say that repudiating one’s role is easy, or cost-free; but it may on occasion not only be morally admirable, but morally obligatory, to do so, even at a high personal cost (in the extreme case, death itself).

What these considerations suggest is that citizenship is first and foremost an activity of human persons, and thus cannot be treated independently from personhood. Citizenship is a relatively important role, but no less than father, mother, friend, teacher, or President, it is unintelligible except as an expression or activity of a human person. Roles are not isolated fragments of activity, to be analyzed piecemeal: they involve the choices and activities of human persons who are responsible or accountable, as persons, for what they do. Thus, in order to make sense of citizenship and its associated excellences, we need to begin by saying something about what it means to be a human person, and what we mean when we speak of the excellence or goodness of a human person.

Needless to say, one could write a book, or several books, on what it means to be a human person, but for the purposes of grounding a concept of citizenship, we need only
focus on some of the most salient traits of persons. I start with persons with their
typical functions developed and in good order, not because other humans are non-
persons, but because even relatively undeveloped humans or humans with their faculties
impaired cannot be understood except in reference to a typically functioning and
developed specimen of the human species. The capacity for personality is what sets
humans above the rest of the animal kingdom: but what does it mean to have a
personality? Whatever else it means, it means at least this much: to be capable of
*reflexively knowing* and *freely and reflexively willing*, both in relation to other impersonal
entities, and in relation to other persons. By *reflexively knowing*, I mean knowing that one
is knowing, and being capable of recognizing knowledge as a good to be sought (or
avoided). By *freely willing*, I do not mean willing in a way that is free from all external
and internal constraint (for that would be a form of anarchy and aimlessness, the paralysis
of all action) but willing in a way that is free from absolute determination to one choice
rather than another. By *reflexively willing*, I mean willing in such a way that one is aware
of one’s own willing and can view it at one remove, so to speak. One may form a view
*about* one’s willing, either good, bad, or indifferent, which may then give rise to what
philosophers call “second-order desires” or desires *about* desires.

We can give this *reflexive knowing* and *free and reflexive willing* some additional
content, for though free, these distinctively human activities are not unconditioned or
entirely indeterminate. On the contrary, human persons are the kinds of beings we hold

---

327 The account that follows is essentially a distillation of the account of personhood given earlier in
§4.2, “The Subject of Integrity.”

328 The *locus classicus* for discussions of second-order desires is Frankfurt, "Freedom of the Will and
the Concept of a Person.”
responsible for their acts: they are capable of perceiving good and evil, and more or less deliberately acting well or ill. Thus, a fuller description of a human person is a being capable of reflexively knowing and freely and reflexively willing good and evil, both in relation to other things, other persons, and himself or herself.

Now, recall Aristotle’s definition of virtue which we already discussed in chapter 4:

It must then be premised that all excellence has a twofold effect on the thing to which it belongs: it not only renders the thing itself good, but it also causes it to perform its function well. For example, the effect of excellence in the eye is that the eye is good and functions well; since having good eyes means having good sight … If therefore this is true of all things, excellence or virtue in a man will be the disposition which renders him a good man and also which will cause him to perform his function well.329

Once we have identified the distinctively human functions as those of freely and reflexively willing and knowing, it follows that virtue (as we saw in §4.5.1 above) is a habitual disposition that perfects the distinctively human functions of knowing and willing, i.e, it is a disposition to know and will well. Although this has a rather formal feel to it, as I suggested in chapter three, there is plenty of broad agreement on what constitutes virtuous or good knowing and willing in a host of situations, and none of our disagreements about virtue should be allowed to conceal this broad agreement.

§6.1.2 Persons and Social Roles

In order to get some grip on the relation between citizenship and the good life, that is, the relation between being a good citizen, and being a good person, we have

begun by taking the concept of person, and the associated concept of personal virtue, as more fundamental than the concept of citizen and the associated concept of civic virtue. This is because being a *person* is something natural, that is, something that *we are by nature*, whether or not we choose it. We cannot repudiate our personhood, since the very act of repudiating it is just an affirmation of it, being a free act of will and intellect. But we can repudiate our citizenship, whether by moving to another country, or condemning our particular traditions or heritage. If we were to ask, why *be* a citizen, or why take the role seriously, we must be able to answer this question in terms of the contribution of the role to a *good* or successful human life. If the role cannot find its place within a successful or well-lived human life, to that extent, the role is morally defective or undesirable.

But to understand more clearly how the citizen role might contribute to a successful human life, how it might be a role a person could embrace in good conscience, we need to say more about (a) what a role is; and (b) what is distinctive about the role of citizen, i.e. what distinguishes it from other human roles. Start with the general concept of a role. During my initial discussion of roles in chapter 4, I defined a social role as

(i) a *function* within society constituted by (ii) one’s social position or relation towards others and (iii) social and institutional norms, goals, and standards of excellence associated with the function, which (iv) shape the practical reasoning of the role-holder and related parties in distinctive ways (chapter 4, §4.6).

Consider, for example, the role of a policeman. How would our definition map onto this role? Well, first, a policeman clearly fulfils a *social function*, namely, that of upholding law and order within society and making citizens more secure in their persons
and properties. This is a shorthand description of the function of a policeman. However, the function is obviously more complex, being shaped by norms, goals, and standards of excellence emanating both ineffably from the social environment (e.g. there is a widespread expectation that the police officer should carry out his duties with courtesy and compassion rather than with cold-hearted mechanical precision); from specific social or legal institutions (e.g. statute law, professional guidelines of the Police Officers’ Association); and from individual police officers as they interpret the role in light of their own judgments (e.g. some police officers may emphasize the importance of building community and thus go beyond the call of duty and strike up a chat with people they run across in the course of duty). Finally, the police officer’s decisions, insofar as they are relevant to the exercise of his police function, in addition to the decisions of related parties (e.g. ordinary citizens, suspects, prisoners), are shaped in distinctive ways by the role he occupies, viz. by the social and institutional norms, goals, and standards of excellence associated by him (and others) with the police function. In other words, insofar as the police officer takes his role seriously and does not see it merely as a theatrical pretence, he will allow its associated norms, goals and standards of excellence to have a direct bearing on action at least insofar as the action somehow involves the performance of his role as an officer. For example, if somebody verbally attacks him in public while on duty, he may consider the dignity of his office and its reputation in the community before responding in kind; whereas it is possible that if verbally attacked off duty, he would be more forthright in responding in kind. Or if he witnesses a crime involving a close friend, he may take steps to distance himself from the friend emotionally in order to prosecute the crime with the full rigor of the law.
Now, compared to the concept of a person, the concept of a role has a notable degree of contingency, in several senses:

(a) Roles do not pertain to human persons as such, but to particular persons or special groups of persons who, whether by design, personal choice, or happenstance, find themselves invited or expected to perform a certain function within society. Roles are distributed on the grounds of natural attributes (such as gender, physical strength, or ethnicity), acquired competence to perform a task (e.g. medical expertise, or knowledge of a field of science), and special ties or bonds towards others, whether permanent (e.g. fatherhood) or transient (e.g. flight instructor); voluntary (e.g. husband or wife) or involuntary (e.g. brother or sister by birth).

(b) Roles are also contingent in the sense that even if we consider a group of persons so positioned that they unquestionably occupy a given role, the role and its interpretation are to a greater or lesser extent conditioned by large-scale social and political variation across space and time: for example, think of the difference between the way people in an “information age” regard the role of doctor and his relation to the patient, versus the way the role of doctor and the doctor-patient relationship was conceived, say, a hundred years ago. Whereas a hundred years ago, it might have been considered impertinent or inappropriate to ask a doctor to disclose the most basic reasons for his medical judgments, it is now considered normal in many situations to do so. Similarly, whereas extended family ties (e.g. cousin, brother) are frequently thought of as incurring an automatic debt of allegiance through thick and thin in Sicily, this is far from the case in the suburbs of California.
(c) Finally, roles are contingent in the sense that their goals and demands are partly open-textured and “essentially contestable”\textsuperscript{330} and thus frequently underdetermine the practical judgments of their occupants. A social role would be nothing but a quixotic individual ritual if it was not constituted, at least in part, by standards, i.e. goals and norms, widely thought to underwrite the role. However, individual persons may, without forfeiting their claim to sanity or reasonableness, interpret an identical role, even within the same society and era, in opposing ways. For example, some argue that judges ought to be maximizers of social utility, while others view them as strict interpreters of the law; some view the function of the teacher as that of conveying technical skills and information in as value-free a mode as possible, while others view a teacher’s function as that of transmitting moral values and “life skills” to his or her students. Or there may be a more localized disagreement, e.g. some believed that Abraham Lincoln was abusing his presidential role by using executive prerogative to suspend \textit{habeus corpus}, while others believed he was remaining faithful to the role; others yet again believed he was temporarily stepping outside the role, but with good reason.

Unless one simply side-steps a role or refuses to occupy it, one must, perforce, interpret its practical requirements in light of one’s own judgments about the goals, norms, and standards of excellence that constitute the role. In other words, the \textit{interpretation} of a role (though certainly some roles a lot more than others) requires creativity and some degree of prudence or practical wisdom rather than the mechanical or rote application of rules—a bit like the performance of a musical symphony which is

clearly much richer than the merely technical performance of the music represented on
the page.

In suggesting that the concept of a role has a high degree of contingency
compared with the concept of a person, I do not mean to suggest that the goals of human
nature can guide our action independently of our roles, relationships, historical context,
and so on. Rather, I mean that we can point to certain features of human beings, such as
their mortality and physical vulnerability; and their intellectual and affective powers,
including their capacity to know, will, and love in a community, that are remarkably
stable over time, and seem to provide a kind of substratum within which roles articulate
themselves, and appear and disappear over time. The basic features of being human are
not something we can change, absent extraordinary and forceful scientific interventions,
and roles, to the extent that they coincide with our humanity (e.g. biological roles such as
parent or sibling), tend to be remarkably stable. So in suggesting that the concept of role
is more contingent than the concept of person, I do not mean to deny that persons’ basic
goals in life are shaped to a significant extent by (more or less contingent) roles, or that
human roles are often given their essential contours by the nature of their inhabitants.

Given that roles are, in part, expressions of human nature, it should come as no
surprise that in spite of the alarming degree of contingency that attaches to many roles—
some roles (e.g. the knight errant) can disappear completely—they are not completely
idiosyncratic social constructs or isolated attitudes: social roles express functions within
human societies, functions that bear some intelligible relation to human well-being. Thus,
the role of political ruler, however it may have changed over the ages, has always served
the purpose of establishing some sort of order within society, bringing some degree of
stability and security into people’s lives, and contributing to their prosperity by providing an environment in which ordinary people can secure a reasonably stable livelihood for themselves and their families. Though there have been tyrannical rulers who undermined such goals or ruled only for their own selfish advantage, they have been recognized as deficient to the extent that they thwarted or sabotaged the social function of the role. Social roles, in other words, cannot be understood independently from human purposes and human well-being. In order to understand the point of a role, we must ask what function it serves in the society in question, and in order to understand its broader practical significance for human life, we must be able to tell a story about how it either furthers the good of human persons, whether it be their external flourishing (e.g. health and wealth) or their all-round excellence; or how it harms some good of human persons, or damages their all-round integrity or well-being. If our goal is to understand or render philosophically intelligible the roles persons occupy, it is not enough to say, “this is the way things are done around here,” or “that was the way things were done back then.” To simply point to the role is to give up explaining it. Roles are susceptible to both a genealogical explanation (why and how did they arise over time?) and an ethical or practical evaluation (how does this role contribute to/detract from the success of individual lives or our life together as a society?).

The location of roles within a wider network of human purposes and goods gives us an important clue as to the limited place a role can have in the practical deliberations of a person. We have just seen that the standards constituting a role call for creativity and judgment in order to be interpreted and applied successfully. However, there is another more far-reaching sense in which roles engage the freedom and practical wisdom of their
occupants: the goals, norms, and standards of excellence that constitute a role are ultimately intelligible only in light of the overarching goals, norms, and standards of excellence of a human life and community. That is to say, roles are not isolated functions: they take their place within a network of functions, including the personal goals and ideals of their occupants. There is nothing sacrosanct or absolute about a social role—it is only as good as the contribution it makes to the life of its occupant, related parties, and the wider social and political community. And so, the judge who applies a facially racist law, the librarian who mistreats customers in the name of official library policies, the father who covers up for his son’s crimes, or the union member who puts hospital patients’ lives in serious jeopardy in order to strike, cannot justify their actions merely by pointing to their role and its requirements, even if (and this is arguable) the role, correctly interpreted, did require such actions. This is not to say that on particular occasions, prima facie unethical behavior may not be justified by the special reasons for action flowing from a particular role. But it is to say that a full ethical account of one’s actions cannot stop at one’s role but must either assume (there are probably cases where it is just obvious) or argue that the role is conducive to the good on the particular occasion in question. For a person is not ethically accountable primarily for the degree to which she is faithful to her roles but for the degree to which she evinces ethical integrity in her interpretation and implementation of them. To put it another way, the challenge of living well is not to pledge allegiance to a series of roles, understood independently from the good life, but to take ownership or responsibility for one’s roles as more or less integral dimensions of a life well-lived. A person is free, even under adverse circumstances, to critically distance himself from the requirements of a role even, in many cases, to the
point of repudiating it entirely (though there are some roles, such as father, which can only be partially repudiated). This if often a costly decision—think of Socrates’s refusal to conform to the popular conception of Athenian citizenship—but it is one that, when made in a responsible and conscientious fashion, is often a tribute to a person’s integrity, and thus merits our praise and admiration.

§6.1.3 Personal Virtue and the Role of Citizen

Now, we are ready to consider the role of citizen in a contemporary liberal democratic society. I will begin by considering the role on its own terms; and then I will consider briefly how it relates to the pursuit of a good human life. First, let us recall our general definition of a social role: A role is a function within society constituted by one’s social position or relation towards others, and social and institutional norms, goals, and standards of excellence associated with the function, which shape the practical reasoning of the role-holder and related parties in distinctive ways. Now, applying this definition to citizenship in a liberal democratic society, we must begin by identifying the function of liberal democratic citizenship. Its function can be described in an abridged form or in a more expansive form. Put most succinctly, the function of liberal democratic citizenship is to contribute to the common good of one’s society and the good of its particular members, understood in light of a broadly liberal democratic ethos. But we can spell out the function of citizenship in greater detail by specifying some of the common and particular goods that call upon citizens’ allegiance in such a regime: These goods include quasi-universal or natural goods such as a stable socio-political order; security of one’s
person and property from attack; and material prosperity\textsuperscript{331}; but also goods that are the peculiar hallmark of a liberal democratic ethos and regime, such as freedom to pursue the good by one’s own lights; freedom of religious belief and practice; freedom to associate with others both in the intimacy of family life and in civil society; and freedom and opportunity to participate and earn a living wage in a market economy; all such freedoms subject to the limits of justice. Citizens of modern liberal democracies are called upon to contribute to the common good of their society, so understood, according to their means: whether by sincerely deliberating about political decisions and taking a hand in public decisions as voters; by proposing themselves for public office; by fostering economic growth and stability in a responsible and long-sighted way; by helping raise children to be responsible citizens; or by building up secondary associations that help instill in citizens a dose of public-spiritedness or concern for the common good, and that educate them in the tools of local if not national politics. Clearly, being a citizen is also defined by more narrowly institutional norms, such as the legal criteria for citizenship; and obedience to the laws of one’s country, to which citizens presumptively owe allegiance, barring special cases of civil disobedience or emergencies (think of the mother who safely breaks a red light to drive her dying child to the hospital).

Finally, there are certain standards of excellence or virtues which are closely associated with the role of citizen insofar as they dispose a citizen to successfully exercise his or her function (i.e. further the common good of his or her society). Some civic virtues that spring to mind are justice, courtesy, civility, tolerance, affability,  

\textsuperscript{331} That these are goods to be furthered by citizens does not imply that they trump all other goods all of the time. For example, social order may become an undesirable good, or not a genuine good at all, if it is a support for severe injustice and political corruption.
conscientiousness, honesty, diligence, and courage. We may call the set of such virtues “civic virtue” for short, as distinct from human virtue simpliciter. But we must exercise caution in interpreting this distinction. It is open to at least two rival interpretations, neither of which seems satisfactory to me. On the first reading, the connection between civic virtue and human virtue is accidental, or fortuitous: civic virtues are derived from the common good of the regime, which, being regime-dependent, has no necessary relation to the actual good of persons; whereas personal virtues are, by definition, excellences of the human person *qua* person, and are thus regime-independent. Thus, civic virtue is artificial and only contingently good; whereas personal virtue, pertaining to human nature, is natural and necessarily good. On a second and quite different reading, civic virtue is just a set of personal virtues viewed from a socio-political perspective. Since the common good of a liberal democracy is a genuine human good, and the civic virtues by definition conduce to the common good, the civic virtues will necessarily form a subclass of the personal virtues.

As to the first interpretation, which rigidly separates civic virtue from personal virtue, or views their connection as accidental at best, this is mistaken in at least two ways: first, it overlooks the fact that social roles are only intelligible in light of the wider purposes they serve; and that the validity of a role such as citizenship (and its attendant excellences) cannot be affirmed independently from its contribution to a flourishing or successful human life. Second, the effort to conceive of human well-being separately

---

332 Aristotle would appear to adhere to such a view, since for him, the quality of civic virtue is determined by the quality of a political regime. Thus, on Aristotle’s view, what is considered civic virtue may often be human *vice*. I must part company with my mentor on this point, since I do not see such a tight and deterministic connection between the quality of a regime and the possibility of a coincidence of civic and human virtue, for reasons I explain below.
from civic virtue and citizenship overlooks the dependency of personal well-being or virtue on contingent features of the society within which it is pursued. After all, conceptions of human virtue and well-being, or of a successful human life, are clearly dependent upon the social and institutional practices of a particular community, and the goods that are internal to those practices. Thus, there is no such thing as a purely “natural” and ahistorical picture of a good human life, or if there is, it is sufficiently abstract that real human beings are compelled to enlarge and enrich it in light of the types of activities, relationships, and commitments their society makes possible. For example, the virtues associated with relatively simple agrarian societies (such as perseverance, diligence, cheerfulness, and loyalty) now have to be supplemented and occasionally modified by more “urbane” virtues such as adaptability to an array of professional skills and tasks; tolerance of diverse lifestyles; and social skills appropriate to a relatively anonymous social environment, such as tact, or the ability to strike a balance between coldness and excessive friendliness to somebody one meets in the subway in a large city.

The second interpretation of the relation between civic virtue and human virtue, which identifies liberal democratic civic virtue unproblematically as a subclass of human virtues, is appealing, but remains unsatisfactory, since it glosses over the messy practical reality of civic virtue and citizenship. On the one hand, if we assume that the purposes, norms, and standards of excellence of a liberal democratic regime can be sincerely and legitimately endorsed as good, then the role of citizen which is constituted by those

333 See MacIntyre, *After Virtue*, 186 ff, for an explanation of the notion of goods internal to practices. MacIntyre stipulates that by a “practice” he means “any coherent and complex form of socially established cooperative human activity through which goods internal to that form of activity are realized in the course of trying to achieve those standards of excellence which are appropriate to, and partially definitive of, that form of activity, with the result that human powers to achieve excellence, and human conceptions of the ends and goods involves, are systematically extended” (187).
purposes, norms, and standards of excellence is a good one, and the virtues prescribed by such a role are genuinely human virtues rather than arbitrary constructs of a political regime. On the other hand, this assumption is problematic, not just because the purposes of a liberal democratic regime can be challenged, but because the purposes of a liberal democratic regime are notoriously fluid and difficult to establish in any publicly convincing way. Of all the social roles one might mention, that of citizenship is notoriously in the category of roles that are “essentially contestable” or the constant subject of debate and negotiation.

Therefore, to posit some pristine role of citizenship serving some pristine set of liberal democratic purposes, in abstraction from the messy reality of role contestation and negotiation, would be to miss one of the most interesting aspects of this role: its inherent fragility and instability. While we may characterize the role in very general terms such as the ones I have suggested above, this does not mean there is a single authoritative interpretation of this general definition that we can appeal to, and then evaluate from an ethical perspective.

In spite of the fluidity and contestability of the role of citizenship, insofar as the general goals, norms, and standards of excellence of a liberal regime are conducive to successful or well-lived human lives both individually and communally, the role of citizenship that is constituted by those goals, norms, and standards of excellence may be interpreted in a way that is consistent with a genuinely virtuous human life. Of course, there is no guarantee that an individual or community will form a conception of civic virtue that is compatible with genuine human virtue, but neither is there any guarantee that an individual or community will form an accurate picture of human virtue.
Indeed, there may be aspects of the dominant ideal (or marginal ideals, for that matter) of personal and/or civic virtue in a community that are deeply defective. It is at times like that that individuals with a rival understanding of personal virtue and its requirements may be forced to repudiate or distance themselves from the community’s notion of citizen virtue, in order to preserve their integrity.

At the limit, there may be cases where a citizen decides upon reflection that the goals, norms, and standards of excellence of a given liberal democratic regime, even stated at a very general level, are incompatible with human goodness. At that point, I believe integrity requires that the role of liberal democratic citizenship be repudiated in whole or in part. However, in many cases a citizen’s dissent is more nuanced: he or she is willing to endorse liberal democratic goals, norms, and standards of excellence at a certain level of generality but interprets the requirements of the common good in a way that conflicts with paradigms of civic virtue that are dominant in his or her community. I think it would be a mischaracterization of this disagreement to say that from this dissenter’s perspective, the role of citizenship, and civic virtue, are simply incompatible with the good life. Instead, it is more accurate to say that, at least from the dissenter’s perspective (and he or she may be mistaken), the dominant conception of citizenship and civic virtue, at least in some respect, is incompatible with the good life. Our dissenter is not repudiating the role of citizenship and its attendant virtues, but offering a dissenting interpretation of them. We must not lose sight of the fact that the disagreement is liable to conceal a significant agreement on the broad contours of the citizen role or function and perhaps even on the broad contours of civic virtue.
To sum up the relation between civic virtue and personal virtue: there is no guarantee that a person’s or community’s understanding of civic virtue will be compatible with personal virtue, anymore than there is a guarantee that a person’s or community’s understanding of personal virtue will be accurate. Nevertheless, to the extent that we deem the most general purposes, norms, and standards of excellence of a liberal democratic regime to be conducive in principle to successful individual and communal human life, there need be no necessary conflict, in principle, between civic virtue and personal virtue. In other words, the role of citizen, assuming the goodness of the general ethos of a liberal democracy, can be construed, at least potentially, as conducive to personal virtue. This “in principle” reconciliation, however, is a constant work in progress, like human integrity itself, and should not blind us to the inevitable tensions that can emerge between influential, socially dominant ideals of civic virtue, and virtue simpliciter. It is enough to think about the dominant atmosphere of intolerance and suspicion towards any form of internal criticism of a country that is at war. Internal critics of the foreign policy of a regime under threat may view themselves as serving the interests of the regime by pointing out its defects; while they may be viewed by others as unpatriotic and even treasonous. This is a conflict not only over how to understand the requirements of personal virtue, but how to understand civic virtue—that portion of personal virtue that contributes to the common good.

---

334 Think of the public indignation and allegations of disloyalty often sparked among Americans by citizens’ criticism of the war in Iraq, especially in its early stages. One does not have to agree with such critiques of the war to view them as good faith acts of citizenship, or what is sometimes called “loyal opposition.”
§6.2 The Aretaic Ideal of Citizenship Applied to Civic Discourse

So far, I have suggested that the citizen role (encompassing civic virtue) in a liberal democracy, though open to constant negotiation and contestation, and vulnerable to construals that put it at loggerheads with genuine personal virtue, can be plausibly construed as not only compatible with personal virtue, but a part of personal virtue that is directed to the common good of a liberal democracy. This is true for two reasons: first, because the common good of a liberal democracy, though susceptible to a range of divergent interpretations by government and citizens, is conducive, in principle, to all-round good human lives, both individually and communally. This is not something I wish to argue in detail, but something I assume to be true, along with most contemporary theorists of citizenship in the Western world. Secondly, because we cannot imagine a good human person who disregards his or her responsibility towards the wider political community, and thus activity of some kind directed towards the common good of one’s society is an essential aspect of human virtue.

In sum, civic virtue is an essential part of personal virtue simpliciter. We might say that civic virtue is the socio-political “face” of personal virtue. Just as personal virtue in general is constituted in part by the timeless attributes of human persons (such as mutual dependency, reflexive knowing and reflexive and free willing) and in part by contingent social, cultural, political and technological developments (think of the virtues distinctive of an agrarian community and those suitable for an urban society), that portion of personal virtue we call civic virtue is constituted in part by the timeless attributes of persons and societies and in part by contingent social, cultural, political and technological developments.
In the next section, I will investigate the conditions that mark civic discourse in modern liberal democratic societies, and the sorts of virtues or character dispositions that allow persons to engage in civic discourse in a liberal democratic regime in a way that is conducive both to the common good and to personal integrity. I hope to show that we can understand the virtues of civic discourse and the resultant ideal of citizenship in a way that is both viable from the standpoint of public order and justice in a liberal democratic regime, and compatible with the full-blooded pursuit of personal excellence or virtue.

But before proceeding with the account, it is worth reiterating its partial or incomplete character. It is selective or partial in the sense that it focuses mostly on the activity of deliberation and discussion among citizens about their common life together; while other activities of citizenship, such as public acts of beneficence and charity, and acts of courage in wartime situations, are not discussed here in any detail. In this respect, my account is no less incomplete than most recent accounts, with this difference: it is more self-consciously incomplete, and to that extent acknowledges the need for a more expansive account.\footnote{335 Most recent accounts of citizenship focus almost exclusively on speech and political deliberation, whether accounts that fall loosely under the umbrella of political liberalism, deliberative democracy, or civic republicanism. See, inter alia, Rawls, \textit{Political Liberalism} ; Gutmann and Thompson, \textit{Democracy and Disagreement} ; Macedo, \textit{Liberal Virtues} ; Benjamin Barber, \textit{Strong Democracy}, 2nd ed. (Berkeley: University of California Press, 2003). A notable exception to this trend is Galston, \textit{Liberal Purposes} .}

The reason I focus on civic speech over and above other activities is first, because I agree with Aristotle that speech is one of the most important distinguishing features of human animals, and one of the most salient \textit{enabling conditions} of politics\footnote{336 See Aristotle’s \textit{Politics} 1, 2, 1253a5-20: “Now, that man is more of a political animal than bees or any other gregarious animals is evident. Nature, as we often say, makes nothing in vain, and man is the only animal who has the gift of speech. And whereas mere voice is but an indication of pleasure or pain…the power of speech is intended to set forth the expedient and inexpedient, and therefore likewise the just and the unjust.”}; and second, because by focusing exclusively on \textit{speech}, as liberal
contractualists such as Rawls do, the comparative merits of an aretaic approach over a contractualist approach to citizenship will be displayed more clearly.

§6.2.1 The Conditions Marking Political Deliberation in Liberal Democratic Societies

As Aristotle remarks in the *Politics*, it is one of the distinctions of the human species that we can deliberate together about what is good and what is just.337 This sets us apart from all other animal species, which are incapable of evaluating and choosing their own actions in a reflective way, and it is this sort of deliberation about the just and the good, ultimately issuing in action, that is one of the distinguishing marks of a political community. Furthermore, since no private society such as a family or club is sufficient to provide the resources necessary for full human development, it is reasonable for persons who aspire to become virtuous or excellent *qua* human beings to acquire the virtues necessary to live in a political community; and not just any political community, but the political community that is available to us *here and now*—at least, provided participation in *this particular community*, with its array of strength and weaknesses, is compatible, in principle, with living a good human life.

Now, if we focus on the preeminent activity of a political community, civic discourse or deliberation about what is good and just for the community, and think about the conditions marking such deliberation within a liberal democracy, we can begin to form a clearer idea about the sorts of virtues required in order to engage successfully in political deliberation in a liberal democratic polity. Some of the conditions shaping such

337 See Ibid.
deliberation are: large populations; extended territory; moral, religious, ethnic, and cultural diversity; a partially centralized and partially free market economy including powerful and influential mass media; and strong, mostly centralized, democratically elected governments.  

**Large populations and extended territory**

Since liberal democracies are composed of large populations (almost always in the millions) and extended territory (for example, it takes a week to drive from the east to the west coast of the United States), the degree to which deliberation can be truly inclusive and face to face is extremely limited. In practice, much political deliberation is mediated either through print and electronic media or representative bodies such as local committees, political parties, and legislatures. A fully inclusive and fully present deliberation of American citizens, or French citizens, or Spanish citizens, about, say, a national law, is inconceivable. Therefore, it is inevitable that a large number of citizens will never sit face to face with a large number of their peers and deliberate with them about the public interest of the community. What sort of political deliberation, then, is open to them? Who can they communicate with about politics? Usually, a limited number of their peers, either informally (say, at home or in a pub), or in a more formal political setting such as a neighborhood association, town hall, or other secondary association. Much less commonly, they can communicate with a large number of their peers by speaking on television or radio, or writing a column or letter to the newspaper, but this

---

338 See Barber, *Strong Democracy* for a detailed discussion of these conditions and some creative ways of adapting democratic procedures to them, including interactive video links in townhall meetings.
form of mass communication is not face-to-face and is much less dialogic than a face-to-face meeting.\(^{339}\)

**Moral, religious, ethnic and cultural diversity**

The populations of modern liberal democratic societies are increasingly heterogeneous in their composition. Populations differ deeply along cultural, religious, racial/ethnic, and moral lines. In spite of the frequent geographic segregation of religiously or culturally distinct groups, the natural mutual affinities that a culturally and religiously homogeneous population can count on as they speak to each other do not exist to anything like the same extent in pluralistic society. Therefore, citizens of liberal democracies are uniquely challenged to communicate in ways that can both engage other moral, religious, cultural and ethnic groups, and in ways that empathize with and take adequate account of their interests and needs.

**A partially centralized and partially free-market economy including the mass media**

The cultural landscape of liberal democracies is shaped pervasively, for better or for worse, by a mixed economy, i.e. an economy that is partly a free market and partly centralized with some government controls and interventions. This environment tends to encourage the relentless acquisition of private wealth, and an “economic” mentality,

\(^{339}\) The dispersed and mediated nature of public deliberation in liberal democracies was brought to my attention by Young’s observation that in a “complex mass society,” we must “[abandon] traces of face-to-face interaction as the model of public discussion, and instead [reinterpret] public debate as mediated among people dispersed in space and time” (Young, *Inclusion and Democracy*, 167). Though any notion of a single, unified and face-to-face public sphere is naïve, I would not go quite so far as to speak of a “subjectless” process of deliberation and I would be hesitant to “abandon” face-to-face interaction as a model of public deliberation, even if we must acknowledge that public deliberation is much more complex than tete-a-tete conversations.
where decisions, including political decisions, are often determined by the “bottom line” or the material cost or benefit to the parties immediately involved. Similarly, the mass media which participate in the market economy may be influenced in their treatment of political issues by factors entirely extraneous to the issues, such as pressure from advertisers to increase sales, or ideological pressures from advertisers to favor certain political perspectives. Political issues may as a consequence either receive scant treatment or be sensationalized or grossly and misleadingly over-simplified.

**Strong, mostly centralized, democratically elected governments**

The concept and reality of the nation-state is a nineteenth and twentieth-century development. Although modern democratic governments are elected, their election is only periodic, and the powers they wield, whether economic, military, technological, or cultural, and the reach and concentration of those powers are without historical precedent. Therefore, the activities of government significantly shape the content and course of civic discourse, whether directly by shaping the terms of the debate in the mass media and passing laws that regulate public discourse, or indirectly by acting on the culture and economy of the nation in a host of subtle and not so subtle ways which inevitably constrain the set of conceptual and practical options available. Furthermore, because modern governments contain a powerful bureaucratic branch that is appointed rather than

---

340 For example, one chat show host was fired shortly after September 11 2000 for suggesting on his program that the suicide bombers who flew the planes into the world trade centers were courageous. Major advertisers threatened to withdraw their support from the show if the show’s host was retained, and the network succumbed to their demands. Whether or not the decision was justified, it illustrates the real influence advertisers can have.
elected, they are not always publicly accountable for their decisions and policies. In short, modern governments exert a powerful influence over civic discourse that is so much a part of the structure of modern societies that it can easily go unnoticed.

§6.2.2 Goals and Virtues of Good Political Deliberation

Now, bearing these four characteristics in mind, what sorts of virtues would a citizen require in order to engage well in civic or political discourse in a liberal democracy? What character dispositions does it take in order to “talk to strangers” about the public interest of our shared political community and to do so constructively or successfully? It seems to me that in addressing this question, we can learn as much from our successes as from our failures. When have citizens been successful at talking to each other, and when have they failed, and why? What is the measure of success or failure in such an enterprise? Although it is not possible to give any formulaic, all-purpose criterion for successful discourse, there are a number of axes of excellence or success in discourse, and none of these standards of excellence can be met without exercising particular human virtues. Below, I offer a list of some of the more salient qualities of a successful political conversation between citizens, and suggest in each case some virtues that would promote the quality in question.

---


342 I borrow this term from Allen, *Talking to Strangers*. 
1. **Practical efficiency and relevance**

2. **Minimization of recourse to violence**

3. **Sensitivity of outcome to interests and concerns of affected parties**

4. **Sensitivity of outcome to ethical and factual critique**

1. **Practical efficiency and relevance of discourse**

Since civic discourse is practical rather than theoretical in orientation, that is, driven mainly by the question, “how should we live together as a political community?” it is normally futile if it does not bear on political action in some way, whether critical or constructive. Political talk that stalls practical progress for the sake of stalling or evading one’s responsibilities (that is, without any good purpose in mind, e.g. the purpose of avoiding a violent confrontation) or fails to advance the parties towards clear practical options in a reasonable time frame, is seriously defective at best, or sometimes even useless. We value practical efficiency and relevance in political discourse, and by this I mean that the conversation addresses the issue(s) at hand; reduces the options on the table to a finite and manageable set; and issues in a practical decision within a reasonable time frame (what is “reasonable” being a contextual and prudential question). What sorts of virtues do parties to a political conversation require in order to attain practical efficiency and relevance in their discourse? Clearly, a certain level of self-discipline or temperance will be required in order to identify the issue at hand and discuss the options surrounding it in a sustained and directed way. Emotional temperance will be required in order to resist the temptation to use the conversation as a pretext for bringing up extraneous issues
(however dear to one’s heart) or picking a peripheral argument with one of the interlocutors whom one happens to find irritating.

Another vital skill for ensuring an efficient outcome is the capacity to listen to one’s interlocutors, not just in the sense of registering the meaning of their sentences, but absorbing the significance, connotations, and tone of what they say, and of coming to terms with its practical implications. This can only be done by enlarging one’s sympathies and imagination to view the world through the eyes of the other, to the best of one’s abilities, and to empathize with the interests of the other. For unless one demonstrably takes on board the interests of others in formulating practical options, dialogue will stall as such narrowly conceived options are unlikely to win the support of other citizens. There may be no name for the disposition to listen empathetically and constructively to one’s interlocutors, but for now let us call it the virtue of active and empathetic listening.

Thirdly, in order to frame the problem accurately and diplomatically, and propose practicable and realistic solutions to it, we require a good dose of prudence or practical wisdom. This virtue enables us to draw sound conclusions about particular actions and situations whose contingency defies any systematic application of a manual of moral rules. Prudence enables us to judge how to act well or successfully in new and more or less unprecedented situations. It does not mean that we disregard general principles or guidelines, but it does mean that the application of general principles to the particulars of a situation or action is shaped in part by those particulars. Many general principles, for example, are defeasible, and the judgment about when they are defeated cannot be made with reference to some further principle of defeasibility, e.g. “only defeated in situation
type X.” For example, supposing a citizen feels obliged, for the most part, to patiently follow the procedures of a town hall meeting. If the meeting takes a nasty turn, e.g. multiple speakers insult one’s ethnicity or religion, she must make a judgment call about whether or not to set aside procedural and social rules, and walk out of the meeting, or speak out of turn. There is no “moral rule book” that she can turn to in order to determine the right thing to do. Her own knowledge, experience, and prudential judgment are what she must rely upon. This virtue cannot be acquired from books (though it may be sharpened or informed by them): it must be acquired through experience and apprenticeship in negotiating practical problems that require a public settlement.

2. Minimization of recourse to violence on either side

Obviously, one of the main advantages of political discourse is that it allows people to air their differences and come to a peaceful settlement of them that, at least ideally, is minimally acceptable to all parties. Thus, even if the settlement involves coercive measures such as laws, if all goes well it depends more on the consent of the parties than the force of violence or coercion. In order for parties to both enter into a peaceful conversation and forgo violent or unilaterally imposed (and thus practically coercive) solutions to their problems, they need to be able to trust the other party to keep his side of the agreement, and not abuse the good will of his interlocutor. In addition, they need to have enough respect or regard for the other party and his interests that they are willing to settle for a compromise that is acceptable to him rather than unilaterally impose their preferred solution, even when imposition is a realistic and legal option (e.g., they are in the majority).
What sorts of virtues make for an environment of mutual trust and voluntary compromise that makes recourse to violence and coercion much less likely? Clearly, honesty, or the disposition to be truthful with others about one’s beliefs, perceptions, and feelings, is essential in order to generate trust. As soon as one is caught in a serious and consequential dishonesty or lie, those who witness it, and in particular those directly affected by it, are much less likely to trust one in future. Closely related to honesty is the virtue of dependability, which is less about accurately expressing one’s intentions and more about following through on one’s expressed intentions and promises. One of the essential conditions for mutual respect and trust in dialogue is the expectation that each party can be counted on to keep his promises and abide by the agreement, in spirit and letter. Again, this trust can only be built up by establishing a track record of dependability, which in turn is made possible by an abiding disposition to be dependable.

The willingness to abstain from violence even in the face of a severe injustice against oneself or one’s “kith and kin” requires at least continence or a substantial degree of rational control of one’s passions, though clearly temperance, where one’s passions are already in line with what is right, is preferable. Contempt, or the treatment of others as inferiors or of no consequence, is identified by philosophers as diverse as Aristotle and Hobbes as one of the vices most destructive of political stability. Therefore, humility, or the disposition to admit one’s own deficiencies as well as merits, and acknowledge the achievements of others even when superior to one’s own, is a virtue essential to political stability and highly conducive to mutual trust and respect.343

---

343 The virtue of humility, or submission to something or somebody greater than oneself and acknowledgement of one’s own neediness, does not appear in Aristotle’s catalogue of the virtues, perhaps because he gives pride of place to magnanimity and the magnanimous man, on his account, hates to be in the debt of anyone. For a good discussion of the concept of magnanimity in Aristotle and how that virtue is
A willingness to make voluntary or unforced compromises depends on a strong sense of mutual trust and respect. We have already spoken of trust. But what about respect? To respect another is to treat another as holding something important in common with oneself. One does not “respect” a blade of grass, except in an exceedingly watered-down sense of the term; perhaps one has a certain sort of respect for the environment and for animals, but this is a very different sense of respect than the respect one bears towards fellow human persons, which stems from something radical or basic one holds in common with them. To respect another is, in some sense, to see the good of the other as one’s own good. Thus, to trample on another human being is somehow to trample upon one’s own person and humanity, to damage oneself at one’s core, to deny one’s humanity so to speak. Now, the identification of one’s own good with that of one’s fellow citizen or human being is a type of friendship. Admittedly, at least in the case of perfect strangers, the identification is much less intense and visceral than, say, the identification by a mother of her own good with that of her child, or the identification by a man of his good with that of his wife; nonetheless, the identification is real and can be viewed, as Danielle Allen (following Aristotle) suggests, as a sort of friendship, which we could call political friendship.344

To have a political friendship with one’s fellow citizens, particularly in a multireligious and multicultural society that extends over a vast territory, cannot be the sort of deep friendship that springs from intimate personal knowledge of another person. But neither can it be a purely instrumental relationship, where one cooperates with the

344 See Allen, Talking to Strangers, 126-139 for an original and insightful discussion of political friendship inspired by Aristotle.

moderated by humility in Aquinas, see Mary M. Keys, "Aquinas and the Challenge of Aristotelian Magnanimity," History of Political Thought 24, no. 1 (2003).
other purely as a means to one’s own ends, and one is prepared to act against the good of
the other insofar as that furthers one’s own ends. To put it in Aristotelian terms, a
political friendship, at least in some respects, is a friendship of virtue, and in other
respects is a friendship of utility. It may not be a friendship of virtue in the sense that
both parties agree completely on what a successful, virtuous life is and wish this for each
other; but it may be a friendship of virtue in the sense that both parties agree on certain
dimensions of a successful, virtuous life, and while disagreeing on substantial points
about the good life, nevertheless wish for each other to excel as human beings, and
acquire the virtues and other resources necessary to live a satisfying and well-rounded
human life. For example, citizens may agree that honesty and diligence are essential
qualities in any self-respecting person, and wish for these qualities to exist in each other;
but they may disagree on various important aspects of the good life, such as the relative
priority of philosophical and political virtues in such a life.

Political friendship is also clearly a friendship of utility, in the sense that one
enters into a relationship with other citizens not just to grow in virtue, or on account of
the other’s goodness, but because some benefit accrues to one from the relationship, a
benefit that is not necessarily intrinsic to the relationship itself.\textsuperscript{345} For example, we
cooperate with other citizens and bear some “friendly feeling” towards them in order to
buy and sell goods, and thus make a living; we speak to other citizens in order to secure
support for policy measures that will favor our business, or neighborhood, or family; or
we bargain with our political adversaries in order to avoid political instability and
violence. Of course, in many cases, political friendship leans heavily towards being

\textsuperscript{345} See NE, 8, iii, 1156a10-30.
purely a friendship of utility. However, in order for genuine respect to characterize political ties, political friendship must be to some extent a friendship of virtue: that is, each party must wish the other well and desire for the other to live well as a human being, and enjoy the conditions required for a successful, virtuous human life. For if political friendship were to exclusively take on the character of a friendship of utility, it would be unstable, since “utility is not permanent but is always changing. Thus when the motive of the friendship is done away, the friendship is dissolved, inasmuch as it existed only for the ends in question.”

Political friendship, the capacity and disposition to bear goodwill towards one’s fellow citizens both for reasons of utility and for the sake of virtue or human excellence, clearly stands out among the virtues of civic discourse, since it provides a fundamental motive for interacting politically with others that extends beyond narrow self-interest or strategic bargaining. One might say that it channels or directs the virtues of discourse, such as honesty and temperance, towards the good of one’s interlocutors. Without a good dose of political friendship, politics would be more like a Hobbesian state of nature than a stable social arrangement; indeed, it is hard to even imagine a citizenry possessing civic virtues such as honesty, humility, justice, and active and empathetic listening, who do not possess the virtue of political friendship to some extent. For it is one of the hallmarks of a virtuous person to wish happiness and virtue to one’s peers and not just for oneself.

To have the virtue of political friendship is to be well-disposed towards one’s fellow citizens, not so much at the level of sentiment as at the level of what one wishes or desires for them. Just as one is willing to make significant compromises with one’s friends in order to accommodate their interests, for the sake of the friendship one bears

346 NE, 7, iii, 1156a20-25.
towards them, so one should be willing to make significant compromises in political discourse to accommodate others’ interests, for the sake of the goodwill one bears towards them, at least if their interests are not nefarious or damaging to the political community.

In fine, in order for civic discourse to issue in non-violent, bilateral agreements, conversation must be marked by mutual trust, respect, and openness to compromise. This is attained when participants bring to the conversation virtues such as honesty, dependability, humility, and political friendship.

3. Sensitivity of outcome to interests and concerns of affected parties

One of the marks of a successful political debate or discussion is that it manages to incorporate within it the interests and concerns of relevant parties, even those who are not physically present. This is deemed a mark of success because we view it as a matter of justice that the interests and concerns of all those involved be given proper consideration rather than marginalized arbitrarily. Modern societies pose a special challenge in this regard, because of their massive scale and the social, cultural and religious heterogeneity of their populations. The problem is partly logistical and partly psychological. Logistically speaking, due both to the scale of political communities and the division of labor within them, it is impractical to involve everyone in common deliberation about the public interest; and it is difficult for relatively small deliberative bodies (in the limit case, two people) to absorb and process the vast amount of information that is required to understand the effects of a policy on a culturally and religiously heterogeneous population. On the other hand, even if these logistical hurdles
were overcome, it is psychologically and morally taxing for any person or group of persons to enlarge their imagination and sensibilities to grasp so many diverse perspectives simultaneously. And unless one already personally knows individuals with the perspectives under consideration, it is hard to see how one would be motivated to process them.

But let us begin with the easier problem: incorporating the interests and concerns of relevant parties who are physically present to the conversation. What sorts of virtues might facilitate such an enlargement of perspective? I have already mentioned one very important one: active and empathetic listening. One may be intelligent and articulate but still be a lousy listener. One cannot possibly enlarge one’s perspective until one has developed the capacity and disposition to listen to and empathize with the perspective of one’s interlocutor. Listening is not just passively registering, but attentively absorbing and processing what one is hearing.

Another virtue not to be overlooked is justice, that is, the disposition to give each his due. Unless one is inclined to give others their due, one easily becomes insensitive to their claims and unmotivated to give their interests due weight in deliberation. Without a just disposition, one easily sees the discourse as a bargaining process in which one optimizes one’s own private interests even where this cancels out the legitimate interests of one’s fellow citizens.

Now, onto the more difficult case of processing the interests and concerns of affected parties who are not physically present to the conversation. Obviously, in any real scenario of political deliberation, whether at home, in the town hall, or in a Congressional or parliamentary debate, this means the vast majority of citizens. There are a variety of
ways in which this absence can be mitigated, e.g. through interest groups that purport to represent a large group of citizens, or through political representatives who travel far and wide to learn about the interests of diverse citizens with a view to securing their votes. However, the gap between those present and those absent to deliberation cannot be closed by institutional devices such as representation and interest groups, and will always pose a challenge to those who are present to deliberation. Some virtues which assist citizens in processing absent perspectives are justice, political friendship, prudence, and enlarged sympathies. Clearly, justice, or the disposition to give each his or her due, is essential in order to motivate deliberators to consider the perspectives and rights of affected parties who are not present. But political friendship, or a feeling of goodwill towards strangers with whom one shares citizenship in common, is just as important as justice, and gives justice a more humane, well-wishing, face. For goodwill does not extend only to rendering justice to one’s neighbors; goodwill extends to desiring a good life for them, on account of their shared humanity in the first place, and in the second place on account of their shared participation in one’s political community. Justice may not engage the imagination and heart in the way political friendship can. Political friendship provides a more reliable and deep-seated motive for extending the imagination to absent strangers than justice does.

Another virtue vital for grasping the perspectives of those who are not present is what I will call enlarged sympathies. This entails the capacity and disposition to imagine the situation and feelings of a person differently-situated by extrapolating from one’s existing experience and knowledge of the human condition. The further removed the other person is, whether in space and time, social situation, or beliefs, the more difficult
such extrapolation becomes. Obviously, someone who lives all her life exposed only to citizens in exactly the same social, cultural, and religious class is unlikely to develop enlarged sympathies to a high degree. An example of the relative absence of this virtue is the difficulty many blacks and whites in the United States often have had in grasping the situation and perspective of each other and how such a perspective informs race relations and race-related policies in the U.S. But other examples of the absence of enlarged sympathies abound: employer-employee standoffs; relations between the poor and the wealthy; and tensions between some religious believers and some proponents of a high “wall of separation” between church and state, to mention a few cases.

Last but not least, prudence or practical know-how, i.e. the ability to apply general principles to hard cases in a way that is sensitive to particulars, is vital in order to effectively take on board diverse citizen perspectives without paralyzing one’s deliberations or experiencing information “overload.” Prudence is required both in order to assess the costs and benefits of taking on board additional perspectives (even in an ideal world, there is insufficient time to process all relevant perspectives) and in order to successfully accommodate diverse perspectives within a single policy. This is not a virtue that can be learnt from books, as I mentioned earlier. It can certainly be fine-tuned by book-knowledge, but it is only the experience of making hard choices time and time again, along with the guidance and example of someone who has acquired the virtue, that can inculcate the virtue of prudence.
4. Sensitivity of outcome to ethical and factual critique

A sound deliberative process will be sensitive to ethical and factual criticism. By this, I mean that it will tend to expose convenient fictions and misconceptions, and will be responsive to political and social reality rather than forming a closed ideological system. In a modern liberal democracy, there are serious obstacles to this kind of ethical and factual sensitivity. Among the most prominent obstacles is undoubtedly the free market economy, for a number of reasons: first, it tends to encourage (if not in principle, then in practice) a pervasively “consumerist” mentality among its members, or modes of reasoning that are instrumentally rational or narrowly oriented towards goals such as material prosperity, comfort, and pleasure. This consumerist mentality may displace deliberation that aims for the public interest or good of the political community with deliberation that aims at maximizing one’s private material acquisitions and comfort. Secondly, the mass media which participates in the free market economy is one of the most influential and sought-after purveyors of public information and argumentation. Yet their public role as purveyor of information, mediator of political debate, and critic of social and political institutions can easily be compromised by market imperatives such as the need to attract and keep advertisers, and the desire to increase or maintain profits in a competitive environment.

Clearly, in a society where the allures of consumerism are so pervasive and where the mass media are often more interested in generating profits than fostering meaningful political debate, constructive political discourse will require some degree of independent thinking and strength of character on the part of citizens. Specifically, it will require temperance in order to subordinate privately-oriented economic motives to the common
good or public interest of the political community; fortitude or courage in order to voice responsibly-held opinions that are not supported by the mass media or public opinion; and honesty, diligence, and patience in order to assess the claims of media organs and politicians in light of rival sources of information or competing perspectives.

§6.2.3 Deliberative Civic Virtue and Integrity

So far, I have considered some qualities of successful political discourse and a range of virtues or character-traits citizens in a liberal democracy would require in order to contribute constructively to it. Our account could stop there, but I am especially interested in how these politically relevant virtues, or civic virtues, relate to virtue simpliciter, that is, to human excellence in general. Recall that human roles, insofar as they are valid or legitimate, must be able to contribute to a good or excellent human life. Just as this applies to the role of teacher, parent, and policeman, it applies equally to the role of citizen in a liberal democracy. Assuming that the goals of a liberal democracy at the most general level are compatible with a good or well-lived human life, the virtues that conduce to those goals and thus enter into the description of the role of citizen in a liberal democracy should be genuine human virtues in spite of their situated and socially and historically informed character. I have considered some of the virtues that conduce to successful political discourse in a liberal democracy, and thus partially constitute the role of citizen: first, what have been traditionally regarded as the four cardinal virtues: justice, temperance, prudence, and courage (or fortitude); then a range of virtues including honesty, patience, diligence, dependability, humility, enlarged sympathies, and active and empathetic listening; and finally a virtue that underwrites the bond of trust and mutual
respect between citizens and is thus of the first importance: political friendship. Now clearly, many of these virtues are required outside the immediate field of political deliberation or conversation, e.g. in the attitudes public officials take towards citizens and in the actions of citizens that occur once they have ceased deliberating. However, this very substantial overlap does not render these virtues vacuous or open-ended: on the contrary, it suggests that our diverse activities as citizens are intimately related to each other, coming from a single agent, and composing, if you will, a single tapestry of virtue (or vice!).

The question that is of special interest to us is, how do these civic virtues connected to political discourse relate to integrity or the pursuit of human excellence in general? It is interesting to note in the first place that many of the virtues required to engage fruitfully in civic discourse are also required in order to persistently seek after the good, e.g. prudence, courage, humility, and temperance. Therefore, unless the other virtues mentioned contradict human excellence (and there is no reason to think they would), the effort to embody the civic virtues in one’s behavior and temperament ought to reinforce, not impede, one’s integrity. Secondly, insofar as the general goals of a liberal democracy, such as human rights, may be construed in a way that is compatible with human flourishing or excellence, allowing one’s actions to be informed by the role of citizenship is to that extent compatible with integrity. And thirdly and finally, the civic virtues are not lived in a vacuum, hermetically sealed off from other virtues or from the overall pursuit of human excellence. A person does not “blindfold” herself to the requirements of human excellence while occupying the role of citizen: on the contrary, she continues to be guided by her best all-things-considered judgment about the good,
taking full cognizance of the duties, responsibilities, and reasons for action that the function of citizenship entails in her situation.

Clearly, in practice negotiating the demands of citizenship and the demands of human virtue is not so neat a task. But this is not because the role of citizen is intrinsically corrupt, but because its dominant legal and social interpretations may contradict human virtue in one way or another. Deciding when, for example, the legal requirements of citizenship such as obedience to the law ought to be set aside, is a difficult problem of conscience, in which the good of social order must be weighed against the benefits of publicly witnessing to the defects of the political order. But this decision is a matter of prudence and integrity: it cannot be decided based purely on some authoritative canon of civic virtue specifying the limits of civil obedience. The role of citizenship and civic virtue are constantly subject to interpretation and re-interpretation, within certain broad constraints. We must not reduce acceptance of the citizen role to blind conformity to the current conventions of citizenship. Otherwise, citizenship is indeed the enemy of integrity.

§6.3 Comparing The Merits of Aretaic and Contractualist Ideals of Citizenship

So far, I have presented a general sketch of the relation between citizenship and ethical life, drawing on the tradition of the virtues exemplified by Aristotle’s *Nicomachean Ethics* and *Politics*; and I have offered an account of the virtues relevant to

---

347 Cf. Aquinas’s discussion of the question “whether human law binds a man in conscience” and specifically whether unjust laws are binding, in the *Summa Theologiae*, I-II, Q 96, a. 4
political speech. In both cases, I suggested that civic virtue—the virtues required in order
to successfully exercise the function of citizenship—is not an arbitrary stipulation of a
political regime, but a necessary condition for a well-rounded, successful human life.
Now, we come to the central claim of this chapter, namely, that an aretaic or virtue-
ethical ideal of citizenship is superior to political liberalism, deliberative democracy, and
agonistic pluralism from a broadly liberal perspective.

Let us begin with what I am calling a broadly liberal perspective: by this, I mean a
philosophical outlook that accepts in principle if not always in practice, the legitimacy
and goodness of the fundamental institutions and practices characteristic of liberal
regimes, such as private property and the free market economy, personal liberty within
the limits of the common good, equality before the law, freedom of religion, freedom of
conscience, freedom of association, rule of law, democratic procedures, and the right to
due process. Now, from this broadly liberal perspective, there are at least three
fundamental values that civic discourse ought in general to instantiate and promote:
respect for persons; political and social stability,\(^{348}\) and the political accommodation of
citizens’ ethical integrity.\(^{349}\)

---

\(^{348}\) Of course, in certain circumstances, it may be desirable or even morally obligatory to provoke
political instability, for example, if the political system is no longer serving the interests of citizens. But I
am assuming that under normal circumstances, stability and peace are desirable from a broadly liberal
perspective.

\(^{349}\) Respect for integrity is intimately connected to freedom of conscience, but to explore that
relationship now would take me too far afield. By treating the accommodation of integrity separately from
the concept of respect, I do not mean to imply that accommodation of integrity is not a manifestation of
respect for persons.
§6.3.1 Respect for Persons

Recall that Rawls’s concept of respect embodies an implausibly stringent reciprocity requirement (the duty of civility) that is liable to alienate many reasonable citizens from the democratic process, and overlooks a broad range of virtues that express mutual respect (§5.2.1.3). Similarly, Iris Young’s concept of respect embodies an implausibly demanding reciprocity requirement in its exclusion of adherence to “unquestioned” assumptions, and overlooks other virtues connected to respect (§5.3.2). Finally, Mouffe’s account of “agonistic pluralism” does not appear to have sufficient moral resources to get any recognizable concept of respect off the ground (§5.3.3).

The aretaic ideal of citizenship, in particular insofar as it speaks to political discourse, embodies a conception of respect that is much richer, truer to life, and more compelling, than the three accounts we have thus far considered. In the first place, the virtues of political discourse lead us to a much more moderate and plausible understanding of reciprocity than Rawls’s. Where Rawls’s principle of reciprocity unreasonably restricts political arguments by insisting that they be plausible to any “reasonable” persons notwithstanding his or her reasonable comprehensive beliefs, the aretaic understanding of reciprocity allows for a high degree of philosophical controversy in political discourse, but insists that citizens display sensitivity to the interests and concerns of their interlocutors, that they acquire the habit of imagining the world from the perspective of their interlocutors, and that they present their arguments rationally, and to the extent feasible, in a way that is intelligible and plausible from the perspective of their fellow citizens.350 As we saw earlier, political friendship, or the disposition to wish

---

350 It seems obvious to me that making a good faith effort to render one’s argument plausible to others by no means guarantees that it will be considered plausible by all – it is perfectly possible, for example, that
others well, is a vital component of respect for persons. Now, since intelligent and free rather than blind and forced obedience to the laws of one’s country is the mark of a mature and virtuous person, a citizen who wishes others well will naturally endeavor to justify common laws to them in ways that are likely to win their rational allegiance. However, a citizen with the disposition of political friendship, while reluctant to override or dispense with the rational consent of other citizens, does not automatically give moral primacy to the principle of consent. Rather, she must make a prudential judgment, in light of other aspects of respect for persons and integrity, what the best course of action is, e.g. whether to defer action until a consensus has obtained, or whether to act in the absence of consent. For example, it may be prudent and respectful to await a high degree of political and social consensus before revolutionizing the taxation system; but if evidence shows that a practice such as smoking is becoming a serious public health hazard, one might judge that it is in the interest of the common good to introduce a smoking ban with only a partial consensus in place. The point is that consent is not a principle that magically trumps all other considerations: rather, it is one important aspect of respect that needs to be considered in citizens’ judgments alongside others.

The second comparative merit of the aretaic understanding of respect is that it is richer and more three-dimensional than Rawls’s, Young’s and Mouffe’s, and consequently, both sociologically and normatively more compelling. The aretaic account fleshes out respect with a range of virtues that many democratic theorists gives little or no attention to. To respect another person, on the aretaic view, is not merely to make a good-

the disagreement runs too deep for any imminent reconciliation to occur, short of a moral or religious conversion. The basic thrust of my critique of Rawlsian reciprocity is that it is asking too much of citizens to present arguments that consistently come across as “reasonable” in spite of fundamental moral, religious, and philosophical disagreements with their interlocutors.
faith effort to justify limitations of autonomy in terms others could accept, but to display a range of virtuous dispositions that evince respect, i.e. a positive regard, for the other person. These dispositions include honesty, truthfulness, dependability, humility, enlarged sympathies, active and empathetic listening; and political friendship. This enlarged account of respect explains why there are many situations where citizens may be ostensibly abiding by the canons of public reason and yet be showing profound disrespect for each other, e.g. by concealing their true motives (dishonesty); failing to acknowledge the achievements and contributions of others (pride or lack of humility); or by listening inattentively or carelessly to one another.

§6.3.2 Political Stability

Increasingly in recent years, we hear of a decline in the quality of public discourse and in the courtesy displayed in citizens’ interactions, from mundane economic exchanges to political debates.\(^{351}\) There is clearly a connection between the moral quality of public discourse and the overall stability of a society. Aristotle singles out speech, and specifically speech about what is just and what is useful, as the capacity that distinguishes human from non-human animals.\(^{352}\) It is our capacity to deliberate about our shared life together that makes politics and political society possible. Insofar as this capacity is stunted or damaged, the very possibility of political society is undermined. Examples of the destabilizing effects of failures to engage in constructive political discourse are the current standoff between the “religious right” and the “secular left” in the U.S.; the

---


\(^{352}\) *Politics* 1, 2, 1252a5-20.
budget standoff between President Bill Clinton and the Republican party in 1999, the many breakdowns and false starts in the Northern Ireland peace process, and the ongoing Israeli-Palestinian conflict. I am not suggesting of course that the low quality of political discourse is the principal cause of these events; but I believe it is one significant cause nonetheless.

Recall that I found Rawls’s, Young’s, and Mouffe’s treatment of political stability unsatisfactory, for a variety of reasons: All three point to the fact of disagreement as a source of instability, but are insufficiently articulate about the specific vices that drive political instability, and the corresponding virtues that might conduce to stability. Rawls and Young offer ideals of civic interaction that are more likely to cause deep division among citizens than constitute a shared ideal of citizenship, while Mouffe’s conception of the conditions of stability seems bereft of ethical moorings.

Pace Rawls and many other democratic theorists, the aretaic ideal assumes that political instability is not caused principally by philosophical and religious disagreements, but by a lack of virtue among citizens, i.e. a failure to habitually act well both towards each other and with respect to the common good of the political community. Here, I focus on the ways in which political discourse can promote or undermine stability, and show how a virtue-ethical account of public discourse is equipped to identify both the sources and the remedies of instability arising from defective political discourse.

Note that I use “democratic theorist” here in a very broad sense, to include most theorists sympathetic to democratic modern democracy. I do not mean by “democratic” that they espouse any particular version of egalitarianism, but I do mean to include theorists such as Mouffe who would reject the label “liberal” yet accept the institutions of liberal democracy.
Political discourse leads to political instability when it falls prey to purely private interests; when it consistently validates gross injustices; when it alienates or excludes substantial groups of citizens; when it fosters mutual distrust and contempt among citizens; and when it conspicuously fails to address pressing social and political problems. Although such failures may be traced, in part, to imperfections and defects in liberal democratic institutions (e.g. debating procedures in representative bodies such as Congress or the British Parliament), we ought not overlook one major cause of defective or destabilizing political discourse: the vices or character defects of citizens. For example, a person with an unjust temperament is unlikely to broaden his horizon beyond his own selfish interests, and thus will tend to gear the public debate towards narrow and private interests to the detriment of the public interest. Similarly, his disregard for the interests and rights of others may often result in the validation of grossly unjust laws; and a person of unjust disposition will care little about whether or not the parties affected by the decision are included in the process or given an adequate voice. Finally, an unjust person, insofar as he bends the political process to his own private interests, will divert the conversation away from pressing social and political problems, many of which may seem unimportant from his individualistic perspective. Similarly, the person with a cramped moral imagination who is not capable of imagining the world from different perspectives or empathizing with other people’s feelings will find it difficult to take on board the diverse interests and perspectives of his or her fellow citizens in deliberation. The proud, contemptuous, dishonest, and un-dependable will cause palpable tensions and animosities to develop among interlocutors, and will seriously undermine rather than foster mutual trust among citizens. For example, rigid adherence to one point of view
even when the evidence tells overwhelmingly against it, suggests an attitude of arrogance and stubbornness that is unlikely to inspire the trust and goodwill of other citizens. Similarly, failure to keep one’s word, especially on a regular basis, makes it increasingly difficult for others to enter into agreements with one.

The virtues I have already considered as important for a healthy political conversation among citizens are antidotes to the all-too-human vices I have just listed as potential sources of political instability: justice, temperance, prudence, courage, honesty, patience, dependability, humility, enlarged sympathies, active and empathetic listening, and political friendship. A just disposition motivates us to consider the interests and rights of others, and the claims of the common good, in our political conversations; while temperance subordinates our immediate desires and attractions to higher desires, such as the desire to live a good life, and to serve the public interest of the political community in our political conversations. Prudence is vital in order to measure the mood and background of our interlocutor and approach him or her or this group or that group in a way that engages rather than alienates them. Courage enables us to address the interests of the community even at great personal cost, and against popular opinion; honesty in conversation, and dependability in action, win the trust, admiration, and goodwill of our fellow citizens; as do patience, humility, enlarged sympathies, active and empathetic listening, and political friendship.

In a difficult and potentially destabilizing standoff between opposing parties, the capacity to imaginatively enter into the worldview of the other, which presupposes both active and empathetic listening, and enlarged sympathies, puts one in a much better position to make a creative offer or gesture of goodwill to break the standoff and thus
restore some measure of political stability. For example, when Al Gore made his concession speech and called off all law-suits in the razor-edge election of George Bush in 2000, he was apparently making a gesture calculated to diffuse the potential resentment of many Republicans, as well as restore stability to the electoral process. The concession could of course be interpreted in purely cynical terms as a last-ditch effort to preserve his own credibility. However, it may also be interpreted, at least in part, as an expression of political friendship or goodwill, and humility or the willingness to acknowledge defeat graciously.

To sum up the superiority of the aretaic ideal of civic discourse to the contractualist one along the axis of stability, it has two points to recommend it: first, it entails a much more moderate, realistic, and plausible understanding of reciprocity, and thus avoids the alienating and destabilizing consequences of “reasonableness” (in the case of Rawls and Young) and ethical groundlessness (in the case of Mouffe); and second, it highlights a much broader range of conditions under which civic discourse is likely to foster political and social stability, and thus alerts us to a broad range of sources of instability and their corresponding remedies. In fine, the aretaic account, while mindful of the relevance of institutional design and social norms to stability, is especially attentive to the character-traits that citizens require in order to deliberate and converse in a way that fosters rather than undermines political stability.

§6.3.3 The Demands of Integrity

Recall that on my understanding, a person lives with integrity to the extent that he approximates three basic conditions: (a) his highest-order desire is to do the good; (b) that
desire is effective and stable enough to dynamically structure and re-structure his moral life (beliefs, actions, evaluations, desires, dispositions) over time; and (c) his beliefs, actions, evaluations, and dispositions are mutually coherent.\textsuperscript{354} The ascription of integrity does not assume perfect coherence but rather is a matter of degrees. However, it seems that a minimum requisite for the ascription is that a person is in the habit of conscientiously structuring his moral life in accordance with his responsible discernment of the good. As I showed in chapter 2, Rawls’s duty of civility imposes a weighty burden upon citizens’ integrity, by placing social and moral restrictions upon their political discourse that their responsibly formed ethical convictions require them to resist. In addition, the principle of reciprocity that governs citizen interactions according to political liberalism, rests on highly controversial moral and epistemological assumptions associated with neo-Kantian or contractualist moral theory. Citizens who hold non-contractualist moral beliefs, such as those associated with virtue ethics and most versions of Christianity, have no good reason to affirm the goodness or ethical soundness of the principle of reciprocity as formulated by Rawls.\textsuperscript{355} Yet Rawls stipulates that observance of the principle of reciprocity is the mark of good citizenship in a liberal democracy, thus branding dissenters bad or morally deficient citizens. I have considered whether such citizens deserve this stigma, and found that in many cases, given their broad support of the liberal democratic ethos, and the reasonableness of their ethical convictions, their dissent from the duty of civility is no good reason to impugn their \textit{bona fides} or civic virtue. The social opprobrium or disapproval that would likely follow upon dissent from

\textsuperscript{354} See §4.5 for my full account of integrity.

\textsuperscript{355} Though some other version, such as “do unto others as you would have them do unto you” seems a plausible candidate.
the duty of civility is thus an unjustified burden upon citizens’ efforts to conscientiously and responsibly pursue a good life, i.e. to live with integrity.

I have also argued that two prominent alternatives to political liberalism, deliberative democracy and agonistic pluralism, do not fare much better in accommodating the integrity of citizens. Young’s conception of “reasonableness” is as philosophically implausible as Rawls’s duty of civility, and liable to alienate many citizens who could be considered responsible and conscientious participants in our political system. Mouffe, on the other hand, while unlikely to accept Young’s ideal of “reasonable” citizenship, lacks the philosophical resources to construct a bridge between the “agonism” of democratic politics on the one hand, and the personal ethical commitments of citizens on the other.

By stark contrast, the aretaic ideal of public discourse, though imposing tangible normative restrictions on citizens’ behavior and interactions, is more friendly to personal integrity than the Rawlsian ideal in at least two respects: first, it imposes a much less question-begging and philosophically partisan burden upon citizens’ integrity than the Rawlsian ideal; and second, it calls upon many of the virtues that are directly constitutive of personal integrity, and thus views civic virtue not as an isolated, role- and regime-specific quality, but as a natural expression of personal virtue in the political sphere.

The aretaic ideal of civic discourse is less philosophically partisan than the Rawlsian ideal insofar as it does not pre-empt the outcome of major philosophical disputes in our culture such as the meaning of autonomy and its relation to other human goods such as friendship, family life, and spiritual renewal. The virtue ethical ideal of discourse recognizes a range of virtues or character traits that permit citizen to engage in
fruitful discourse, including justice, temperance, prudence, courage, honesty, dependability, humility, and political friendship. These virtues, though often in short supply among citizens, resonate across a broad range of secular and religious philosophies, in no small part because their importance is a matter of shared experience rather than abstract stipulation. The relatively ecumenical nature of the virtues of political discourse makes them a good candidate for a public ideal of citizenship that is much more accommodating of integrity than contractualist ideals such as the three we have considered.\(^{356}\)

A second notable virtue of the aretaic ideal from the perspective of integrity is that, being richer and more wide-ranging than the contractualist ideal, it calls upon many of the virtues that we have already found to be directly constitutive of personal integrity or the responsible pursuit of the good. Thus, it allows us to view civic virtue, in this particular case in the realm of public discourse, as an expression or overflow of personal virtue rather than a merely regime- or role- specific requirement. Some of the virtues of civic discourse which correspond directly to the virtues of integrity (i.e. the virtues required in order to persistently order and re-order one’s life in accordance with one’s responsibly-formed judgments about what is good) are temperance, prudence, courage, honesty, and humility. These virtues are required both in order to live a good life—

\(^{356}\) It is important to note that the aretaic ideal, while a lot more ecumenical than Rawls’s, neither pretends, nor aspires, to be philosophically “neutral” or accessible to any “reasonable person.” Some degree of philosophical controversy and contestation is inevitable. The advantage of being relatively ecumenical is that the ideal can provide a viable starting-point for many public conversations and can realistically inform the public philosophy of democratic nations.
however one’s political regime or situation may be—and in order to engage competently and responsibly in civic discourse in a liberal democracy.\textsuperscript{357}

\section*{§6.4 Objections and Replies}

In order to round out my account of the virtues of political discourse, I would like to respond to a few objections that I would imagine a skeptical—or for that matter, sympathetic—reader raising.

1. The \textit{aretaic ideal of citizenship is so complex and general that it is difficult for citizens to draw specific conclusions from it for their behavior in the political sphere.}

This is essentially a version of the objection that virtue ethics is too vague in its practical implications to be action-guiding in the way, say, deontological or utilitarian ethics are.\textsuperscript{358} This objections calls for a two-part response: first, as Rosalind Hursthouse quite eloquently argues, all ethical theories, including deontological ethics, appeal to formal practical premises which can only be interpreted and applied to concrete situations by “filling in” the details or interpreting the premises according to our moral intuitions and/or experience. So, when Kant insists that we act as self-legislating rational beings, this premise is no more action-guiding than “be virtuous” or “be prudent”: its interpretation requires reliance on our moral knowledge, which extends far beyond the immediate content of the premise. Thus, though there are always going to be hard cases

\textsuperscript{357} It would be surprising if these same virtues were not essential to healthy civic discourse in \textit{any} political regime. However, for the purposes of this essay, I restrict my account to the conditions of liberal democratic societies.

\textsuperscript{358} This objection is put most forcefully in \textit{ask David Solomon}. See \textit{Solomon}, \textit{?} \textit{##} for a response on behalf of virtue ethics.
that call for difficult prudential judgments, in general we can know what honesty calls for in many situations with little difficulty, and formulate general, though often defeasible, principles to express our understanding of the practical implications of this virtue: for example, do not deliberately deceive your fellow citizens about your true intentions; tell the truth to the best of your ability; do not exploit legal procedures to gain an unfair advantage over your opponent. And so on for the other virtues.⁵⁵⁹

The second response to the action-guiding objection is that moral judgment almost always involves an element of prudence. Though it would save us a lot of work if we could reduce an ideal of citizenship to some simple rules of behavior, to do so would be to blind oneself to the complex and multi-faceted nature of moral judgment.⁵⁶⁰ Abandoning a mechanical, casuistic understanding of ethics need not lead to moral anarchy or some crude form of intuitionism. For judgment itself, however difficult or complex, presupposes standards of judgment. The aretaic ideal does not propose a single, overriding standard of judgment, but a complex standard, composed of the various civic virtues and rules associated with them.⁵⁶¹ Applying this standard is a matter of prudence. On many occasions this is an easy judgment call, while on others, it is more difficult.

---


⁵⁶¹ These are what Hursthouse calls the “V rules.” For example, “Be honest” would be a relatively high-level V-rule, while “Do not cheat,” “Do not lie,” and “Tell the truth when you owe it to someone,” would be applications of that higher-level rule, if you will. This is not to suggest that rules, by themselves, can be correctly applied by anyone, however vicious or silly or immature. The correct application of rules in particular cases requires a range of virtues (depending on the situation), pre-eminently the virtue of prudence or practical wisdom. These may be learnt by study, experience, practice, and apprenticeship, but not by study or intellectual learning by itself.
2. It is difficult to discern any clear implications of the aretaic ideal of citizenship for institutional design, legal reform, or public policy.

The first obvious response to this objection is that an ideal of citizenship is distinct from ideals of institutional design, legal reform, and public policy. Ideals of citizenship hold up ethical standards against which citizens should measure their behavior, not principles of institutional design, legal reform, or public policy. The task I have set myself here is to develop an aretaic ideal of citizenship, and that task is of independent interest within liberal and democratic theory. That said, there is no reason to assume that fleshing out the institutional and legal implications of the aretaic ideal of citizenship is an impossible task. For example, if honesty is the mark of good citizenship, then it would seem unwise to design political institutions that, wittingly or unwittingly, incentivize or reward dishonest behavior. Furthermore, students of liberal democracy are increasingly attentive to the relevance of civic virtue to public policy and institutional design. 362

3. The aretaic conception of reciprocity abolishes Rawls’s clear-cut criterion of legitimacy (the principle of reciprocity) but puts nothing clear in its place. Thus, it robs the political process of any clear criterion of legitimacy.

As far as providing a clear-cut criterion of legitimacy is concerned, any simple formula for legitimacy seems doomed to failure, since legitimacy is a complex notion that depends on a variety of factors, from procedural propriety to the substantive justice of the

regime or policy under consideration. For example, in certain cases we are willing to trade off the criterion of consent against the substantive justice of the policy under consideration. Think of a case such as school segregation in the United States. Many would contend that even if *Brown v. Board of Education* (1954) overrode the consent of reasonable persons, it was still legitimate because it was overturning a severe and longstanding injustice. This is not to deny that consent and procedural justice are central to any plausible understanding of political legitimacy in a liberal democracy. But it is to deny that consent and procedural justice can be considered as necessary or essential conditions for legitimacy in every case. Civic virtues such as justice, temperance, courage, enlarged sympathies and political friendship obviously contribute to the legitimacy of a regime to the extent that they make for a citizenry that is watchful towards the interests of others and capable of working together patiently towards just political solutions. But the liberal fascination with simple formulae for legitimacy, from Mill’s harm principle to Rawls’s principle of reciprocity, while understandable, invariably reduces and oversimplifies complex and prudential political problems. It is time we reconciled ourselves with making judgments about political legitimacy and justice that are informed by a variety of criteria which, on pain of dismissing the role of judgment or prudence entirely, must be central considerations in political judgments rather than quasi-mathematical formulae. The search for a single, simple formula for legitimacy is about as likely to meet with success as the search for the Holy Grail.
4. The aretaic ideal of citizenship is unrealistic since it expects citizens to exercise an impressive range of virtues in their mutual interactions in a world in which, we all know, virtue is in short supply.

There is no denying that I am proposing a demanding and difficult-to-attain ideal of citizenship. Few if any citizens will ever attain it in its fullness. However, like any ideal (including Rawls’s), the aretaic ideal of citizenship is first and foremost a benchmark for good citizenship, not a description of the actual interactions of citizens in liberal democracies. If citizens frequently fall short of the ideal, then at least we have some measure against which to understand their failure (or indeed to grasp the measure of their success).

Secondly, I would be the first to concede that if the aretaic ideal was utterly superhuman, presenting a standard so far removed from human life that only an angel could even approximate it, then it would be difficult to know how it could serve as an ideal for flesh-and-blood human beings. However, this ideal of citizenship, though challenging and in some respects utopian, is realistically utopian, to use John Rawls’s phrase. That is to say, it presents an ideal that citizens can hope to approximate to a substantial degree. Of course, hardly anybody ever instantiates the fullness of civic virtue, in all its dimensions, just as hardly anyone achieves the fullness of personal virtue. But history and our own personal experience undoubtedly throw up examples of citizens who instantiate to a significant degree at least some of the civic virtues that constitute the aretaic ideal. We all know, or know of, citizens who have been exemplary in their courage, or humility, or justice, or fair-mindedness, or honesty, or prudence. Granted, few manage to exemplify all the civic virtues in their full splendor, but the fully rounded-
out picture of civic virtue is a useful measure of good citizenship nonetheless, pointing as it does to the qualities that must occur, at least in a dispersed form, in order for public discourse and politics to function relatively smoothly and justly in a liberal democracy. If civic life is enriched by the diverse qualities of all citizens, then we can at least hope that the respective strengths of different citizens might complement each other and even that citizens might inspire one another by manifesting different virtues in their behavior.

Of course, to paint an ideal of civic virtue is very different from claiming that civic virtue is alive and well in our society. Indeed, it would appear that civic virtue is in many respects on the decline.\textsuperscript{363} However, insofar as my account posits a tight connection between civic virtue on the one hand, and respect and political stability on the other, it offers a potential explanation for the shortage of respect and socio-political stability in modern liberal democracies such as those of Europe and North America. To point to the need for increased virtue in modern democratic societies is not of course to prescribe a practical remedy to our problems: but it is at least to set a goal which we must keep in mind as we deliberate about how to address the social and political problems confronting modern democratic societies. The aim of this dissertation is not to prescribe specific policies or political solutions, but to convince citizens and policymakers that they cannot afford to ignore the relevance of civic virtue as an essential condition for mutual respect, political stability, and the reconciliation of personal excellence with public life.

\textsuperscript{363} See Carter’s interesting reflections on the decline in civility and related virtues, in Carter, \textit{Civility}. William Galston goes so far as to observe what he “can only regard as the relentless tribalization and barbarization of American life” (Galston, \textit{Liberal Purposes}, 6).
5. Many of the virtues specified by the aretaic ideal are just Rawls’s virtues of reasonableness and reciprocity under another name: e.g. sensitivity to the interests of others, justice, empathetic listening, etc. How is the aretaic ideal truly distinct from reasonableness?

It is true that there is substantial overlap between Rawls’s virtues of reasonableness and the virtues specified by the aretaic ideal. And I consider that not a weakness but a strength of my account: it gives us many features of Rawls’s account without the drawbacks (e.g. the deleterious implications of reasonableness for citizens’ integrity and the stability of the regime and the contrived narrowness of Rawls’s canon of the civic virtues). But it does substantially add to Rawls’s account, in several respects: first, it adds to Rawls’s canon of virtues virtues he hardly gives any treatment of, such as temperance, humility, and prudence, which make a crucial contribution to the health of civic life; and second, it reinterprets the cooperative virtues so that they must be understood to function not in isolation from the other virtues, but to the extent possible, in conjunction with them. For example, a tolerant person without prudence or courage would become pusillanimous and might well “tolerate” practices or behavior that ought to be confronted and challenged if not legally, then socially, such as racial and ethnic slurs.
6. For all the difficulties that afflict the duty of civility, the alternative is equally if not more destabilizing. For example, when people are permitted to invoke controversial philosophical and religious reasons that fall outside “public reason,” many citizens will reject the outcome of the process as an illegitimate imposition grounded in this or that publicly inaccessible comprehensive doctrine.

The fear of the destabilizing influence of comprehensive doctrines trades on an extreme interpretation of the inclusion of comprehensive doctrines, as well as a misunderstanding of the major sources of instability in political life. First, we do not have to choose between the total exclusion of dependence on comprehensive doctrines on the one hand (the duty of civility) and the unrestrained reliance upon them on the other. The aretaic understanding of reciprocity, while it does not rule out reliance on comprehensive doctrines, does expect citizens to exercise prudence\(^{364}\) and consideration in the sorts of arguments they advance for policies and laws. Insofar as they wish other citizens well or bear them political friendship, they will endeavor to secure their rational consent and collaboration by offering them arguments they are likely to find plausible and intelligible. For example, when a Christian argues with a secular humanist, it would seem wise to begin with the ethical values and claims that unite the two parties, such as freedom and human virtue, rather than those that divide them, such as God’s Fatherhood of the human race. And this need not be based merely on the need to secure political support to pass a proposal: there is an intrinsic (though not absolute or unconditional) value to winning the rational allegiance of one’s fellow citizens for one’s political proposals, given the

\(^{364}\) By prudence, I must reiterate that I do not mean the modern sense of instrumental rationality, but the Aristotelian sense of practical wisdom, reasoning well about particulars.
importance of rational self-direction or autonomy in any flourishing human life. On the other hand, there may be occasions (e.g. cases of life and death such as the permissibility of assisted suicide) where the two parties have reached an impasse and the disagreement can no longer be rendered intelligible without recourse to the resources of their respective comprehensive doctrines. Yet again, there may be occasions where the philosophical difference between the two parties is so great that the only realistic option is to present one’s arguments openly and honestly, knowing in advance with near certainty (though rarely with complete certainty) that they will be rejected outright by one’s interlocutor. In short, though the absolute prohibition of dependence on comprehensive doctrines is unwarranted, the aretaic account counsels prudence and consideration for the outlook of one’s interlocutor in advancing arguments and in employing the resources of one’s comprehensive doctrine.

7. The aretaic ideal seems to implicitly stigmatize or exclude forms of expression that are not “rational” and “civil” or courteous. Though it rejects Rawls’s formulation of the duty of civility, it does not reject civility as a virtue, and it does emphasize the virtues that conduce to political stability. This ideal does not appear to have room for protest and disruptive political participation, which is often both morally necessary and admirable.

---

In response to this objection concerning the stigmatization of political protest and “uncivil” forms of political communication, I would begin by pointing out that political stability is generally and for the most part considered a good thing in developed democratic societies, while social and political instability is usually considered a necessary evil rather than something to be pursued for its own sake. Thus, it seems legitimate to focus on the regime-stabilizing effects of the civic virtues, assuming that a certain level of stability and order in a regime is desirable, and assuming that the regime is not so corrupt that the principal form of political engagement available to its critics is protest and disruption.

However, nothing in my account so far excludes a prominent role for protest and disruption in civic engagement. On the contrary, to the extent that protest is a responsible, thoughtful, and respectful response to injustice or political corruption, virtue ethics has the resources to describe it as an act of virtue compatible with civility rightly understood. Unlike Rawls’s duty of civility, the duty to be civil, honest, compassionate, sincere, humble, and so on, understood from an aretaic perspective, can only be interpreted in light of the requirements of other virtues, in particular those virtues pertinent to a situation. One such virtue is justice, or the disposition to render each (including the political community) his due. Another is prudence, or the disposition to discern and act according to what is best for all concerned (including the actor) in particular situations. If civility is reduced to consistent mannerliness or “niceness,” it becomes a pseudo-virtue with little inherent value, just as dependability does if it is reduced to unquestioning faithfulness to all of one’s commitments. The point of a virtue is to act well, not to

366 This would be an example of what Young calls internal exclusionary rules within a deliberative process that is ostensibly open to everyone. See Young, Inclusion and Democracy, 55.
blindly follow a single pattern of behavior. But this can only be achieved if we have other virtues (e.g. justice) to correct the distortions or excesses of individual dispositions such as civility whose general orientation is positive or virtuous but which may at times become misdirected or vicious.

Besides allowing that apparent “incivility” can be an expression of virtue compatible with civility rightly understood, virtue ethics guards against considering protest against injustice to be decisive evidence of virtue or good character. Just as civility is only truly a virtue when corrected and shaped by a complex of virtuous dispositions such as justice and courage, so the courage to protest is only truly a virtue when it is animated by a just disposition, respect for one’s fellow citizens, and a concern for the common good rather than merely one’s own selfish gains. Paradoxical as it may seem, even radical protest must be undertaken in a responsible and respectful manner, with the right dispositions. If not, its moral quality is impugned.

**Conclusion**

In chapters 2 and 3 I explicated the political liberal ideal of citizenship and exposed its internal incoherence and its hostility to personal integrity. In order to strengthen the case against political liberalism from integrity, I then proceeded to outline a relatively robust concept of personal integrity and defended its political and moral viability against the most salient liberal objections (chapter 4). In chapter 5, I offered a deeper and broader diagnosis of the failure of political liberalism, setting out more concisely the criteria of success for an ideal of citizenship, and arguing that political liberalism and much contemporary political theory fail to meet these criteria partly on
account of its contractualist foundations. In this final chapter I have sketched a positive alternative to Rawlsian citizenship, namely, an aretaic ideal of citizenship that takes its ethical bearings both from the liberal tradition of representative democracy and from the classical tradition of the virtues. The sketch is necessarily selective and incomplete, focusing for the most part, not on the whole gamut of civic activities, but on the vitally important activity of political discourse. Nonetheless, insofar as a genuinely virtue-ethical ideal grapples more successfully or convincingly than political liberalism and other contractualist accounts with the problem of political deliberation in a liberal democracy, it must be considered a potential rival to contemporary contractualist theories of politics—not just in the sphere of political deliberation, but in its treatment of political problems in general.

I began my constructive account of citizenship by outlining in very general terms the concept of citizenship as a role, and its relation to ethical life broadly understood, bringing two paired terms into relation with one another: the nature of a person and personal virtue on the one hand, and the role of a citizen and civic virtue on the other (§6.1). I then presented some of the most important civic virtues relevant to political discourse (§6.2), and argued that the aretaic ideal of political discourse is demonstrably superior to contractualist ideals, insofar as it offers a richer and more realistic concept of respect for persons; proposes a more wide-ranging and fruitful diagnosis and treatment of the sources of political instability; and is considerably more accommodating of the demands of integrity (§6.3). I concluded the chapter by responding to seven of the most important objections against my virtue-ethical ideal of citizenship (§6.4).
The aim of this chapter has not been to lay down a definitive or complete model of citizenship, but to sketch, in broad outline, part of an aretaic ideal of citizenship, and compare it with the contractualist ideal along the axes of respect, stability, and integrity. My hope is not that the reader will agree with every detail of my account, but that she will at least be convinced by my analysis that the classical tradition of the virtues cannot be easily dismissed as a serious rival to contractualist political theory in the quest for a viable ethical basis for modern citizenship. Perhaps the reader will even be convinced that the classical tradition of the virtues holds out more promise than political liberalism (and other liberalisms insofar as they share its limitations) for a theoretically cogent and practically viable ideal of citizenship in a liberal democratic regime.
CONCLUSION

The starting-point for this investigation is a burning question that most recent accounts of modern citizenship are insufficiently attentive to: is it possible to wholeheartedly pursue a good life and simultaneously remain in good standing as a citizen in liberal democratic polities such as those of Europe and North America? Can citizens remain faithful to their deepest ethical and religious values and commitments while pledging allegiance not only to their laws but to public norms of civic conduct and deliberation? This question takes on a special urgency at a time when the tone of civic discourse seems increasingly shrill and the concept of a public philosophy seems to be brought into question by the deep constitutional, philosophical and religious questions that divide us. It is tempting to believe that we are faced with a stark choice: either “privatize” our personal moral and religious values and take on the garb of public reason upon entering public life, thus assuring some semblance of common ground with our fellow citizens; or reject the values of public reason, pledge fidelity to our personal ethical values, and settle for a public life ruled by democratic procedures, strategic bargaining, and majority coalitions rather than principled ethical and moral deliberation.

Yet both of these choices are extremely costly, and were we forced to choose between them, our situation would be tragic indeed. What both choices have in common is resignation to some form of moral disintegration: in one case, disintegration of one’s most cherished ethical commitments upon entering the political sphere; while in the
other, the disintegration of any public ethical or moral foundation for political life. The principal goal of this dissertation is to chart a course between these two choices, one that vindicates the hope, even if it does not guarantee the achievement, of some reconciliation between ethical and political life. Spurred on by this aspiration, this dissertation has grappled with a series of questions all of which relate, in one way or another, to the quest for a more ethically integrated interpretation of public life and citizenship: (i) what do we actually mean by integrity or the unity of a human life? What does it mean to be undivided or to have one’s ethical commitments intact? (ii) What is the blindspot that prevents contemporary theories of citizenship from engaging the question of integrity at the deepest level? (iii) What are the social, political, and ethical ramifications of the failure to rise to the challenge of integrity? And (iv) can we hope for an account of citizenship that might rise to the challenge of integrity without turning its back either on public norms of conduct or the possibility of a free society?

In order to get to the bottom of the relation between integrity and modern citizenship, I have undertaken a careful study of the ideal of citizenship proposed by John Rawls in *Political Liberalism* (1993). I have taken Rawls’s political liberalism and its implicit ideal of citizenship as sufficiently paradigmatic of modern liberal and contractualist approaches to citizenship to warrant very close scrutiny, not simply for its own sake, but in the hope that its limitations might provide some clue about how to diagnose the ethical limitations of many other recent accounts. This hope did not prove to be in vain: In my critique of Rawls’s ideal of citizenship, I was led back to the inadequacy of its most basic normative foundation, namely the principle of reciprocity. I found that insistence upon the primacy of consent in political deliberation could only be
explained by Rawls’s embrace of a vision of social and political life as a contract among equals, a vision in which morality itself was somehow constituted by the imagined consent of reasonable and rational adults.

This broader diagnosis of political liberalism’s failure to convincingly confront the challenge of integrity (chapter 3), along with my more refined interpretation of the concept of integrity (chapter 4), provided a framework within which many of Rawls’s contemporaries could also be critiqued. My critique of political liberalism led me to frame the hypothesis that the contractualist moral imagination prevents many theorists of citizenship from perceiving the power of virtue ethics to speak to three central problems of modern political order: (1) developing a plausible conception of respect; (2) accommodating the ethical integrity of diverse citizens within a single polity; and (3) providing the conditions necessary for a stable political order. In chapter 5, I tested this hypothesis against two of Rawls’s rivals from different strands of political theory: Iris Young (deliberative democracy) and Chantal Mouffe (agonistic pluralism), and found that they too, in spite of their differentiation from Rawls, labored under the limiting assumptions of contractualism.

Finally, confronted with the limitations of contractualism, in chapter 6 I proposed a rival account of citizenship rooted in the virtues—what I call the “aretaic” account—and showed how it fared better along the axes of respect, stability, and integrity, than its contractualist peers. What makes the aretaic account especially appealing from the perspective of my project is that it holds out the prospect of a normative foundation for politics that it is both responsive to the special demands of the role of citizenship, and respects the integrity of a person who seeks to live an all-round good life. I hasten to add
that I am not proposing any quick-fix solution to the tensions between politics and the ethical commitments of citizens. Nonetheless, I am proposing a framework within which those tensions can be illuminated in a way that creates some hope of a *rapprochement*—however delicate or fragile—between personal and political ideals of conduct.

As we reach the end of this project, it seems fitting to point out that it is only just the beginning of another one: the project of developing an ideal of citizenship for modern life rooted in the tradition of the virtues that has enjoyed a renaissance among analytic philosophers and political theorists in the past two decades or so. Though there is much talk of virtue among political theorists these days, the focus tends to be on the virtues narrowly tied to political deliberation and toleration, such as Rawlsian reasonableness and Young’s virtues of inclusiveness. The most pressing challenge facing political philosophers sympathetic to virtue ethics is to bring a wider and richer account of virtue to bear upon political problems, one that embraces tolerance and cooperativeness to be sure, but does not shun humility, patience, magnanimity, and friendship. The account of civic virtue offered here is a narrowly focused and preliminary application of virtue ethics to the sphere of political deliberation, and is offered not as a systematic treatise on civic virtue, but as a modest primer for a political philosophy that broadens and deepens its attention to the virtues.

Finally, one of the more interesting questions my preliminary inroads into civic virtue naturally give rise to is, “what is the relation between this moral ideal of citizenship and public policy, whether in the sphere of education, family law, commerce, or the arts?” Though my account explicitly confines itself to the ethical ideal of citizenship rather than its policy implications, the ideal cannot be implemented or
realized without some clear sense of the sorts of policies and laws that might further or thwart it, and the role of competing institutions, whether government, civil society, or family, in promoting civic virtue. The aretaic account of citizenship offered here, though abstracted from policy questions, is intended to provide basic elements of a sound ethical vision of citizenship that can provide a solid foundation for the discussion of practical policy proposals. After all, it is surely of the first importance that we get straight what sort of society we are aiming to achieve before jumping headlong into policy discussions. If this dissertation does nothing but make a significant contribution to that foundational task, it will have served its goal.


