

## Rawls and the Kantian Ethos\*

Nicholas Tampio

Hamilton College

*John Rawls had a life-long interest in Kant. To provide a new perspective on Rawls's political thinking, to illuminate Kant's legacy for political theory, and to contribute to current debates about the Enlightenment, I track how Rawls interprets and transforms Kant's legacy. In this essay, I show how Rawls reconceptualizes four key Kantian activities: the identification of the problem, the engagement with common sense, the construction of principles, and the authentication of principles. I defend Rawls from the charge—made by Allan Bloom, Michael Sandel, and Allen Wood, among others—that Rawls fundamentally misunderstands or misuses Kant. The basis of my defense is that Rawls considers a critical intellectual sensibility (or ethos), rather than a specific doctrine (e.g., the categorical imperative), as the most valuable component of Kant's legacy. Polity (2007) 39, 79–102. doi:10.1057/palgrave.polity.2300044*

Keywords Rawls; Kant; Enlightenment; ethos

*Nicholas Tampio teaches political theory at Hamilton College. He researches the legacy of the Enlightenment in contemporary political theory. He can be reached at nicholas.tampio@gmail.com.*

### Introduction

John Rawls, it is well known, was a life-long student of Kant.<sup>1</sup> One of the most famous and controversial sections of *A Theory of Justice* (1971) is entitled, “The

---

\*This essay grew out of seminars on Kant and Rawls at Johns Hopkins University with William E. Connolly, Richard E. Flathman, Eckart Förster, and J. B. Schneewind. I also thank Jane Bennett, Harry Gould, George Klosko, Patrick Peel, Sophie Rosenfeld, Andrew Ross, John Tambornino, Lars Tønder, Stephen K. White, the anonymous reviewers at *Polity*, and audiences at the Johns Hopkins University Graduate Student Colloquium, the University of Virginia Political Theory Colloquium, the New York State Political Science Association Conference, and the Western Political Science Association Conference.

1. The following abbreviations will be used for Rawls's books and essays:

IM “The Independence of Moral Theory,” in *Collected Papers*, ed. Samuel Freeman (Cambridge: Harvard University Press, [1975] 1999).  
IP “The Idea of Public Reason Revisited” (1997), in *Collected Papers*.  
JF *Justice as Fairness: A Restatement*, ed. Erin Kelly (Cambridge: Harvard University Press, 2001).  
KC “Kantian Constructivism in Moral Theory” (1980), in *Collected Papers*.  
LH *Lectures on the History of Moral Philosophy*, ed. Barbara Herman (Cambridge: Harvard University Press, 2000).

Kantian Interpretation of Justice as Fairness” (Section 40). The 1980 essay that signals the shift from early Rawls to late Rawls is “Kantian Constructivism in Moral Theory.” In 1989, Rawls wrote an essay on Kant, “Themes in Kant’s Moral Philosophy,” that elucidates Rawls’s own political theory. In *Political Liberalism* (1996) and *Justice as Fairness: A Restatement* (2001), Rawls no longer identifies his theory as Kantian, yet Kant is a regular interlocutor in these pages. Rawls details his debt to Kant in the recently published *Lectures on the History of Moral Philosophy* (2000). Rawls places Kant at the center of his account of the history of moral philosophy and examines Hume, Leibniz, and Hegel primarily to clarify Kant. In a reminiscence that serves as the foreword to *Lectures*, Rawls explains how he taught Kant to his classes at Harvard:

My efforts were centered on trying to understand him so as to be able to describe his ideas to the students. Sometimes I would discuss well-known objections to his moral doctrine, such as those of Schiller and Hegel, Schopenhauer and Mill. Going over these is instructive and clarifies Kant’s view. Yet I never felt satisfied with the understanding I achieved of Kant’s doctrine as a whole. . . . All the great figures . . . lie to some degree beyond us, no matter how hard we try to master their thought. With Kant this distance often seems to me somehow much greater. Like great composers and great artists—Mozart and Beethoven, Poussin and Turner—they are beyond envy (LH, xvii).

The aim of this essay is to explicate Rawls’s relationship to Kant and the larger intellectual-political tradition of the Enlightenment.

Rawls’s critics often focus on his interpretation of Kant. Straussians, civic republicans, and agonistic democrats are among those who challenge Rawls’s reading of Kant to undermine Rawls’s political theory. In an early essay on *A Theory of Justice*, Allan Bloom argued that Rawls cannot synthesize the utilitarian social contract tradition and austere Kantian morality. For Bloom, Rawls seeks to retain Kant’s ideas of freedom and rationality without adopting the principle of universality. Rawls desires the “glow of Kantian moral nobility” without demanding the heroic sacrifices of Kant’s moral principle. For Bloom, Rawls misuses Kant to create a nihilistic political theory.<sup>2</sup>

Michael Sandel also suggests that Rawls’s flawed reading of Kant indicates a deeper problem with Rawls’s political theory. In his early and late phases,

---

PL *Political Liberalism* (New York: Columbia University Press, 1996).

TJ *A Theory of Justice: Revised Edition* (Cambridge: Harvard University Press, 1999).

TK “Themes in Kant’s Moral Philosophy” (1989), in *Collected Papers*.

2. Allan Bloom, “Justice: John Rawls versus The Tradition of Political Philosophy,” *American Political Science Review* 69 (1975): 656.

Rawls envisions a “deontology with a Humean face,” that is, a political theory that grounds Kant’s idealistic moral principles in an empiricist framework. For Sandel, a coherent political theory may have one feature or the other but not both. Early Rawls entangles himself in the transcendental metaphysics that he hopes to avoid; late Rawls situates himself in the contemporary political milieu but loses the ability to articulate principles that transcend that milieu. Rawls’s political theory “either fails as deontology or recreates in the original position the disembodied subject it resolves to avoid.”<sup>3</sup> Early Rawls goes himself on the second horn of this dilemma, late Rawls on the first. Both versions of Rawls’s theory, for Sandel, are philosophically incoherent.

Bonnie Honig agrees with Bloom and Sandel that Rawls misunderstands Kant. For Honig, Rawls misconstrues Kant’s project and his own if he believes that contextualizing his theory of justice makes it substantially more political than Kant’s. Political liberalism, according to Honig, fails to politicize citizens or teach them about the individual and collective workings of power. Both Kant and Rawls act as if politics is “a sign of failure, a Fall.”<sup>4</sup> The major difference between Rawls and Kant, for Honig, is that Kant realizes that his theory of justice displaces politics while Rawls does not.

Kant scholars offer a surprisingly tepid defense of Rawls’s reading of Kant. On the one hand, Kant scholars regularly acknowledge that Rawls elevates Kant’s visibility in contemporary moral and political philosophy.<sup>5</sup> Several prominent Kant scholars, including Paul Guyer, Christine M. Korsgaard, Onora O’Neill, and Susan Neimann, wrote their dissertations under Rawls’s supervision at Harvard. Many Kant scholars respect Rawls’s accomplishments as a philosopher and as a teacher. On the other hand, Kant scholars often contest Rawls’s interpretation of Kant. Larry Krasnoff asks in a recent article in *Kant-Studien*, “How Kantian is Constructivism?” According to Rawls, constructivism provides a complex conception of the person, a procedure of construction, and the ability to manufacture shared principles. Krasnoff explains that Kant’s categorical imperative concerns concrete individuals rather than abstract persons, specific tests of maxims rather than generalized procedures, and a shared method rather than common principles. “Hence Kant’s moral philosophy does not conform to *any* of the three features of constructivism

---

3. Michael Sandel, *Liberalism and the Limits of Justice: Second Edition* (New York: Cambridge University Press, 1998), 14.

4. Bonnie Honig, *Political Theory and the Displacement of Politics* (Ithaca: Cornell University Press, 1993), 136.

5. Allen W. Wood, *Kant’s Ethical Thought* (New York: Cambridge University Press, 1999), 337; J. B. Schneewind, “Autonomy, Obligation, and Virtue: An Overview of Kant’s Moral Philosophy,” in *The Cambridge Companion to Kant*, ed. Paul Guyer (New York: Cambridge University Press, 1992), 333–34.

specified by Rawls.”<sup>6</sup> For Krasnoff, Rawls ought to leave the field of Kant scholarship to the specialists.

This essay defends Rawls’s approach to reading Kant. Initially, I analyze Rawls’s interpretation and reworking of four key Kantian activities: the identification of the problem, the engagement with common sense, the construction of principles, and the authentication of principles. Then, I show how Rawls justifies his refashioning of Kant’s legacy on Kantian grounds. For Rawls, the critical disposition impels one to address the problems of one’s time and create ideas and principles accordingly. Rawls embodies a Kantian ethos when he transforms Kant’s legacy.

Rawls’s reading of Kant, I contend, impacts contemporary debates about the Enlightenment. Theorists on the Left and the Right have recently attempted to pin down once and for all the ideals of the Enlightenment.<sup>7</sup> Kantians often enter these debates by positing that Kant’s doctrines define the Enlightenment. Allen W. Wood, General Editor of *The Cambridge Edition of the Works of Immanuel Kant*, states:

Kant’s ethical thought . . . exercises such a strong and continuing influence on us that replacing commonly accepted ideas about it with more accurate and less oversimplified ones might help to transform our conception of our own history and of ourselves as heirs of the Enlightenment.<sup>8</sup>

Heirs of the Enlightenment, according to Wood, ought to protect Kant’s doctrines against corruption. Rawls, on this view, is a chief culprit in Kant’s misrepresentation and simplification.<sup>9</sup> I maintain, on the contrary, that Rawls displays how to reinvigorate the Kantian tradition and the Enlightenment in general. The motto of the Enlightenment, according to Kant, is “have the courage to use your *own* understanding!”<sup>10</sup> Rawls exercises this courage when he creates political theories appropriate to late-modernity. Rawls shows that the most valuable legacy of the Enlightenment is a dynamic way of thinking rather than a static set of doctrines.

6. Larry Krasnoff, “How Kantian is Constructivism?,” *Kant-Studien* 90 (1999): 401. See also Wood, *Kant’s Ethical Thought*, 374–75, and Onora O’Neill, *Constructions of Reason: Explorations of Kant’s Practical Philosophy* (New York: Cambridge University Press, 1989), chapter 11, and “Constructivism in Rawls and Kant,” in *The Cambridge Companion to Rawls* (New York: Cambridge University Press, 2002), chapter 9.

7. See, respectively, Stephen Eric Bronner, *Reclaiming the Enlightenment: Toward a Politics of Radical Engagement* (New York: Columbia University Press, 2004), and Gertrude Himmelfarb, *The Roads to Modernity: The British, French, and American Enlightenments* (New York: Knopf, 2004).

8. Wood, *Kant’s Ethical Thought*, 14.

9. See note 6, above.

10. Immanuel Kant, “What is Enlightenment?” in *Practical Philosophy*, ed. Mary J. Gregor (New York: Cambridge University Press, 1996), 8:37, 18. Kant citations provide the pagination from *Kants Gesammelte Schriften* and the pagination from *The Cambridge Edition of the Works of Immanuel Kant*.

## Identifying the Problem

The problems of political liberalism are not the problems of Enlightenment liberalism (PL, xl). Rawls, a pragmatist, disputes the idea that philosophy specifies “a fixed family of problems with agreed criteria for deciding when they are resolved” (LH, 17). The intellectual and political problems of the twenty-first century differ from the problems of the eighteenth century. Rawls therefore doubts Kant’s principles may directly address our concerns.<sup>11</sup> Yet Rawls draws from Kant and Enlightenment liberalism to frame the problems of political liberalism.

Three historical developments, according to Rawls, shaped the problem of modern moral philosophy (PL, xx-xxx; LH, 1–17). First, the Reformation fragmented the religious unity of the Middle Ages. Medieval Christianity was an authoritarian religion. It taught that there was no salvation outside of the single church and that apostasy was a sin and a crime. It articulated a creed that each individual had to believe without question. The Reformation challenged the hegemony of the Catholic Church. Luther and Calvin wanted to replace one authoritative and salvationist religion with another, but their actions had unintended consequences. The “fact of religious division” in sixteenth and seventeenth centuries Europe motivated philosophers to envision alternatives to the theocratic state.

Second, the birth of modern science helped dissolve the medieval worldview. For certain thinkers (such as Hume), Newtonian physics signaled the deathknell of Christian morality; for others (such as Leibniz and Kant), the relationship between religion and science became more complicated. In either case, the relationship between reason and faith changed as a result of the discoveries of Copernicus, Newton, and Galileo.

Third, the birth of the modern state transformed the background conditions of modern philosophy. Philosophers needed to explain how the modern administrative state enriches and threatens human freedom. Despite their disagreements, Enlightenment philosophers concurred that the problem of modern practical philosophy was to establish a comprehensive secular moral doctrine (PL, xx).

Kant’s more precise problem was to articulate an alternative to the prevalent doctrines of his day: Hume’s utilitarianism and Leibniz’s rational intuitionism. Kant agreed with Hume on several points (PL, xxix). Both believed that the moral order arises from the confluence of human nature and social life. Both thought

---

11. It is instructive here to contrast Rawls’s aims with Wood’s. In his book, Wood intends “to present Kant’s ethical thought in a new light, in order that it might be easier to understand and more relevant to ethical questions of today” (333). For Wood, Kant’s principles remain valid even if his judgments, or his application of those principles, do not. For Rawls, Kant’s principles themselves need to be rethought after the philosophic critiques of Hegel and Dewey and historical developments such as the Civil War and the New Deal.

that moral awareness is available to every reasonable and conscientious person. And both postulated that human beings do not need the commands of God or the state to be moral. Kant criticizes Hume, however, for failing to appreciate our nature as rational beings. Hume's utilitarianism provides the rudiments of a theory of justice. By thinking from the perspective of a judicious spectator, we can imagine what administration of social resources will maximize the satisfaction of our individual systems of desire (TJ, 29). Hume's utilitarianism helps satisfy the goods we desire as (primarily) sensible beings. For Kant, however, utilitarianism permits injustice by permitting the good of the collective to have priority over the rights of the person (TJ, 23). Kant thinks social welfare may never override civil liberties. The right, for Kant, frames the good and restricts it absolutely. Kant seeks a theory of justice that prioritizes our nature as rational, rather than merely sensible, beings, that is, that applies pure practical reason to the conditions of human life (TK, 510).

The second moral doctrine that Kant addresses—Leibniz's rational intuitionism—seems to solve the problem of utilitarianism (KC, 343–46; TK, 510–16; PL, 90–98). Kant and Leibniz agree that basic moral concepts exist independently of the natural world and may be grasped solely by rational intuition. Kant disagrees with Leibniz on the source of moral concepts. For Leibniz, the source is the transcendent hierarchy of perfections; for Kant, it is human practical reason. Kant states rational intuitionism is “heteronomous.” What (on Rawls's reading) is the force of Kant's charge of heteronomy? A heteronomous doctrine is “compatible with a variety of contents for the first principles of a moral conception” (TK, 511). There is no way to resolve the dispute between rational intuitionists such as Plato and Leibniz over the definition of rational principles. This leads to what Rawls, in *A Theory of Justice*, calls “the priority problem” (TJ, 37). Up to now, no rational intuitionist has systematized the plurality of first principles discovered by pure reason. This is a serious problem given rational intuitionism's thesis that “the agreement in judgment which is so essential for an effective public conception of justice is founded on the recognition of self-evident truths about good reasons” (KC, 344). Kant aims to create a theory of justice that is “autonomous,” that is, that only relies upon “our conception of ourselves as reasonable and rational persons” (TK, 512). This theory must also satisfy the “content condition,” that is, provide sufficient guidance for moral deliberation (LH, 163).

We may now understand (from Rawls's perspective) the general and specific problems that Kant addresses in his moral philosophy. The general problem of Enlightenment Liberalism is to replace Christianity with a secular comprehensive morality (IP, 611). Kant's specific problem is to create a moral doctrine that avoids the dangers of utilitarianism and rational intuitionism. Rawls describes Kant's specific aims as follows (PL, 99–101). First, Kant formulates a moral doctrine that gives the idea of autonomy a regulative role for all of life. For Kant, the moral

concepts that we discover in pure practical reason should apply equally to our public and nonpublic lives. Second, and related, Kant thinks that moral philosophy should aspire to constitutive autonomy. That is, Kant maintains that there is an objective order of values (contra utilitarianism), but that this order can only be constituted by the activity of human reason (contra rational intuitionism). Third, Kant formulates his conceptions of the person and society on the basis of transcendental idealism (rather than Hume's empiricism or Leibniz's transcendental realism). Fourth, Kant presents his entire moral philosophy as a defense of reasonable faith. This reasonable faith is, for Kant, the expression of each individual's sound human reason. In short, Kant aims to put forth a comprehensive moral doctrine grounded on a metaphysical framework.

Rawls, as noted above, thinks the problems of practical philosophy have changed in the last two hundred years. First, philosophers have become wary of formulating comprehensive secular moral doctrines. One explanation is that we continue to experience the aftermath of the Reformation (PL, xxvi). In the past three centuries, Europe has witnessed "the successful and peaceful practice of toleration in societies with liberal institutions" (PL, xxvii). Philosophers see that social unity and concord does not require agreement on a comprehensive moral doctrine. Philosophers also appreciate the theoretical and practical "burdens of judgment." That is, philosophers now recognize that "different conceptions of the world can reasonably be elaborated from different standpoints," and that "our individual and associative points of view, intellectual affinities, and affective attachments, are too diverse, especially in a free society, to enable those doctrines to serve as the basis of lasting and reasoned political agreement" (PL, 54–58). We still need to think about the problem of justice, according to Rawls, but we also need to address the question of tolerance. The general problem for political liberalism, therefore, is: "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)?

If Rawls's general problem is how to formulate a political conception of justice, his specific problem is how to distinguish it from two prominent contemporary alternatives: the utilitarianism of J. S. Mill, J. C. Harsanyi, and R. B. Brandt (TJ, 20n), and the rational intuitionism of Henry Sidgwick, G. E. Moore, and W. D. Ross (TJ, 30n; KC, 343). Although each of these thinkers postdates Kant, the parallel with Kant's distinction between empiricist and rationalist heteronomous moral doctrines is clear. On the question of foundations, Rawls is closer to empiricism than to rational intuitionism. From *A Theory of Justice* on, Rawls has recast Kantian ideas and procedures "within the scope of an empirical theory" (TJ, 227). Despite his theoretical agreement with many of the presuppositions of empiricism, Rawls opposes the prevalent form that empiricism has taken in contemporary ethics, that is, utilitarianism. The problem with utilitarianism, for

Rawls, is that it cannot account for our “common sense convictions” that all rational human beings possess certain basic rights and liberties (TJ, 24–25). According to Rawls, any satisfactory political doctrine in the contemporary world must model our belief in the priority of the right to the good (TJ, 25).

The second main contemporary moral doctrine that Rawls considers—rational intuitionism—seems better equipped to construct a conception of justice that respects the rights of the person. According to Rawls, however, rational intuitionism has not gotten any closer to solving the priority problem (TJ, 36–40). Rawls emphasizes that Kant’s relevance for constructing a political conception of justice decreases at precisely those moments when he embraces rational intuitionism.<sup>12</sup> For a contemporary theory of justice to be morally defensible, as well as pragmatically viable, it must serve different purposes than the one Kant envisioned (PL, 99–107). First, it must articulate a political conception of justice rather than a comprehensive moral view. We need to find a political conception that is both moral and able to command the assent of an overlapping consensus of reasonable comprehensible doctrines. This requires a conception of justice that does not require everyone to endorse the ideals of constitutive autonomy or believe that the only appropriate conceptions of the person or society for political morality originate from transcendental idealism. Rawls is not indifferent or skeptical to Kant’s doctrine of morality, but he thinks that it is undesirable (and impossible) to expect widespread assent to it. Finally, Rawls believes that we need a new understanding of reasonable faith.

Political liberalism, once again, does not address the problems of Enlightenment liberalism. Rawls sees limits to how much contemporary political philosophy can learn from authors writing in the comparatively slow-paced, homogeneous world of eighteenth-century Europe. Yet Rawls appropriates themes from Kant as he formulates the problem of political liberalism (PL, 93–101). First, Rawls believes the content of political justice may be represented as the outcome of a procedure of construction (though both the content and the structure should be described differently). Second, Rawls thinks the procedure of construction should be based on practical, not theoretical, reason (though Rawls has a different conception of pure practical reason than Kant does). Third, Rawls argues that the relevant conception of the person for moral philosophy must be relatively complex (though this conception now draws on the fundamental ideas in the public political culture rather than on transcendental idealism). Finally, Rawls thinks that moral philosophy should assume the task of defending reasonable faith (though this is now faith in the “possibility of a just constitutional

---

12. For a useful discussion of the development of Rawls’s relationship to Kantian metaphysics, see Erika Flickshuh, *Kant and Modern Political Philosophy* (New York: Cambridge University Press, 2000), 20–30.

regime”). Rawls, in short, reactivates Enlightenment liberalism by reconceptualizing its problems.

## Engaging Common Sense

Early and late Rawls believes that practical philosophy must engage common sense, that is, the opinions that most people hold about morality and justice.<sup>13</sup> *Lectures on the History of Moral Philosophy* reveal that Rawls views this Socratic activity through Kantian lenses. Rawls also thinks that common sense has changed over the past two centuries and that it is futile to repeat Kant's conversation with common sense. For Rawls, Kantians ought to update Kant's engagement with common sense rather than merely clarify Kant's dialogues in the eighteenth century.<sup>14</sup>

Rawls describes the strategy of Part I of Kant's *Groundwork of the Metaphysics of Morals* (1784) as follows:

The aim of the argument, which seems valid, is to find the supreme principle of morality (the moral law). It starts from ordinary commonsense moral knowledge and moves to philosophical knowledge by elucidating the underlying principle found in our everyday judgments about the moral worth of actions (LH, 153).

Rawls highlights several aspects of Kant's strategy in Part I. First, Kant engages “ordinary commonsense moral knowledge” to protect it. Kant does not introduce a new principle of morality to common sense but discloses its principle so that it is not confused by bad moral philosophy (LH, 158). What, though, does Kant mean by “ordinary commonsense moral knowledge”? Ordinary commonsense moral knowledge is our unreflective awareness of “what is right and what is wrong” (LH, 148). For Kant, ordinary commonsense moral knowledge expresses itself most clearly in Christian concepts. That is why Kant analyzes the concepts

---

13. In his first published essay, “Outline for a Decision Procedure of Ethics” (1951), Rawls aims to determine a method that will clarify the principles underlying “the many dictates of commonsense morality,” in *Collected Papers*, ed. Samuel Freeman (Cambridge: Harvard University Press, 1999), 7. This essay is rewritten as Section 9 of TJ (“Some Remarks about Moral Theory”), pages 8–9 of PL, and pages 29–32 of JF.

14. Recent attempts to clarify Kant's engagement with common sense include Roger J. Sullivan, *An Introduction to Kant's Ethics* (New York: Cambridge University Press, 1994), chapter 2, and Paul Guyer, *Kant on Freedom, Law, and Happiness* (New York: Cambridge University Press, 2000), chapter 5. Sullivan and Guyer present Kant's argument in Part I of the *Groundwork of the Metaphysics of Morals* as if it has timeless validity. See Immanuel Kant, *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, ed. Mary J. Gregor (New York: Cambridge University Press, 1996). For Rawls, this manner of reading Kant misses the religious background of Kant's thought and, therefore, the limited application of Kant's argument to contemporary political circumstances.

of the good will and the world as holy in Part I of the *Groundwork*. For Rawls, the “religious, even Pietist, aspects of Kant’s moral philosophy seem obvious” (LH, 161).

Second, Kant engages common sense to reform as well as to defend it. Kant explains that though “there is something splendid about innocence . . . what is bad about it, in turn, is that it cannot protect itself very well and is easily seduced.”<sup>15</sup> Common sense may succumb to a “natural dialectic” that permits individuals to prioritize their sensible inclinations over their moral duties. Sound (Christian) common sense, in other words, can degenerate into unhealthy (hedonistic) common sense. That is why Kant ends Part I of the *Groundwork* by leaving the sphere of common sense to enter the field of pure moral philosophy. For Kant, the engagement with common sense is a prolegomena to the metaphysical, and authoritative, argument of Part II of the *Groundwork*. “Taken alone, much that Kant says [in Part I of the *Groundwork*] is misleading and can be understood only in light of what comes later” (LH, 157). For Kant, there is a chasm between common sense and pure moral philosophy.

Rawls engages common sense in a way that both parallels and diverges from Kant. Rawls states his aim in Kantian terms:

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 (KC, 306; see also TJ, 18; PL, 8).

Rawls modifies Kant’s engagement with common sense in several ways. First, Rawls updates the description of healthy common sense. Today, healthy common sense is democratic rather than sectarian. Rawls finds this common sense in texts and traditions other than the Bible and Christendom. “In a democratic society there is a tradition of democratic thought, the content of which is at least familiar and intelligible to the educated common sense of citizens generally. Society’s main institutions, and their accepted forms of interpretation, are seen as a fund of implicitly shared ideas and principles” (PL, 14). How does Rawls help democratic common sense? Primarily by adjudicating the debate over how to balance the competing goods of liberty and equality. According to Rawls, democratic common sense is currently divided between those elements (influenced by Locke) that give priority to the liberties of civic life and those elements (influenced by Rousseau) that give priority to the political liberties (KC,

---

15. Kant, *Groundwork*, 4:404–05, 59.

307; PL 4–5). Rawls offers his theory of justice—justice as fairness—as a tool for democrats to formulate a suitable public understanding of freedom and equality.

Rawls also agrees with Kant that philosophy ought to criticize common sense:

It is important to keep in mind the subordinate place of common sense norms. Doing this is sometimes difficult because they are familiar from everyday life and therefore they are likely to have a prominence in our thinking that their derivative status does not justify. None of these precepts can be plausibly raised to a first principle. Each has presumably arisen in answer to a relevant feature connected with certain particular institutions, this feature being but one among many and these institutions of a special kind. Adopting one of them as a first principle is sure to lead to the neglect of other things that should be taken into account. And if all or many precepts are treated as first principles, there is no gain in systematic clarity. Common sense precepts are at the wrong level of generality. In order to find suitable first principles one must step behind them (TJ, 270–71).

Why does Rawls think philosophy ought to “step behind” common sense? The main reason is that “there is a tendency for common sense to suppose that income and wealth, and the good things in life generally, should be distributed according to moral desert” (TJ, 273). There are traces here of Kant’s critique of hedonism in Part I of the *Groundwork*. For both Kant and Rawls, human beings have a natural tendency to rationalize their desires even if those desires are immoral or unfair. Rawls observes that wealthy and fortunate individuals often justify their good wealth and fortune on the basis of desert. Conservative common sense asserts platitudes such as, “Justice is happiness according to virtue” (TJ, 273). Rawls contends, however, “the initial endowment of natural assets and the contingencies of their growth and nurture in early life are arbitrary from a moral point of view” (TJ, 27). Rawls aims to strengthen democratic common sense and weaken conservative common sense, to take, in sum, an active role in “the conflict between liberalism and conservatism at the present time” (PL, 22).

Rawls differs from Kant as well in how he conceives the relationship between moral philosophy and common sense. For Kant, as noted above, the sphere of pure moral philosophy transcends the field of common sense. For Kant, the sphere of pure moral philosophy has authority over the field of common sense because it alone is grounded on pure practical reason and not on a “marvelous mixture” of sentiments and desires.<sup>16</sup> Rawls—influenced by Hume, Hegel, and Dewey—aims to overcome these dualisms in Kant’s moral philosophy (TJ, 226; KC, 304). Rawls thinks practical philosophy must go back and forth between

---

16. Kant, *Groundwork*, 4:410, 64.

analyzing common sense and constructing philosophical concepts (TJ, 18–19; PL, 8–9). To think about justice, according to Rawls, we ought to “work from both ends,” that is, alternate excavating common sense and considering philosophical ideas. We seek reflective equilibrium in which our judgments and principles coincide. We aspire to “wide” reflective equilibrium in which we have considered a broad array of common sense convictions and philosophical theories. Reflective equilibrium, for Rawls, is a provisional accomplishment: we will always encounter new developments in academia and the world. Interminability, for Rawls, is the price a Kantian philosophy in an empiricist framework must pay.

### Constructing Principles

One of Kant’s great contributions to moral and political philosophy, according to Rawls, is his notion of constructivism (KC). “Kant is the historical source of the idea that reason, both theoretical and practical, is self-originating and self-authenticating” (PL, 100). In *Political Liberalism*, Rawls distinguishes his version of constructivism (“political constructivism”) from Kant’s (“moral constructivism”). Yet Rawls’s political constructivism replicates the major ideas and steps of Kant’s moral constructivism even as it amends them in significant ways.

The leading idea of Kantian constructivism, according to Rawls, is “to establish a suitable connection between a particular conception of the person and first principles of justice, by means of a procedure of construction” (KC, 304). Kant’s moral constructivism takes the following form. First, it defines the relevant conception of the person for moral philosophy as the “rational being as such” (*vernünftige Wesen überhaupt*). According to Kant, this is the conception implicit in individual moral consciousness (KC, 340). It is also the conception provided by transcendental idealism (PL, 100). Kant represents the person as *vernünftig* to characterize two of its moral powers (PL, 48–54; TK, 503–04). One moral power is empirical practical reason. This is the power to use means-end reasoning to attain desired objects. A rational being as such has the capacity to formulate general precepts (all under the category of *the* hypothetical imperative) to help it realize its conception of happiness. The other moral power is pure practical reason. This is the power to take an interest in the moral law. A reasonable being as such has the capacity to formulate laws (all under the category of *the* categorical imperative) that enable it to act morally regardless of circumstances.

The next stage of moral constructivism is to construct principles by following a certain procedure. This procedure is not itself constructed but rather laid out (TK, 513–14). The task of a Kantian is to outline, in procedural form, all of the conditions on reasoning that Kant seems to rely on in his moral philosophy. If we

incorporate all of these conditions into a procedure, and we follow that procedure correctly, then we have *prima facie* assurance that our resulting moral judgments will be reasonable. The basis of Kant's moral constructivism, according to Rawls, is the conception of ourselves that we elicit from ordinary moral experience. Kant's moral constructivism mirrors, and thereby helps us see more clearly, the ideas implicit in our self-conception as reasonable and rational persons.

One way to construct the content of Kant's moral doctrine, according to Rawls, is to consider how Kant lays out six conceptions of the good (TK, 506–10). By following this ordered sequence, we gain insight into how Kant envisions the application of pure practical reason to the conditions of human life. In particular, we see how Kant seeks to give appropriate weight to our desire for the good (by showing how all six conceptions of the good connect to empirical practical reason) but also to assign absolute priority to our interest in the right (by showing how the higher conceptions of the good are framed by pure practical reason). The first conception of the good is given by unrestricted empirical practical reason. This is our conception, here and now, of happiness. The only principle that we act on to achieve this conception of the good is that we ought to do what we must to maximize our happiness (*the* hypothetical imperative). The second conception of the good is that of the fulfillment of "true human needs." True human needs, according to Kant, are the general ends all human beings have regardless of particular ends. The first and second conceptions of the good mirror our rational desire for happiness (though the second lays the ground for interpersonal comparison: a prerequisite for moral thinking). Beginning with the third, Kant shows how higher conceptions of the good incorporate an independent conception of right. The third conception of the good is that of the fulfillment in everyday life of "permissible ends." To determine whether the ends of actions are permissible, we test them by the procedure of the categorical imperative. The "CI-procedure," in brief, enables us to test whether our subjective principle of action, if generalized and made into a law of nature, would result in a "perturbed social world" that we would affirm and in which we would still intend to act from our maxim. This third conception of the good mirrors our reasonable and rational belief that everybody ought to be free to pursue their conception of the good as long as it does not violate the moral law. The fourth conception of the good that Kant describes is that of the "good will." This is the ideal for which each individual human being should strive: it is the pure expression of our nature as reasonable and rational beings. A human being with a good will always acts from, and not merely in accordance with, the moral law. Finally, the last two conceptions of the good specify the ideal world that we, as rational and reasonable beings, ought to try to produce. The fifth conception of the good is that of the "realm of ends." The realm of ends is simply the social world that

would result if all reasonable and rational beings acted in accordance with the moral law. The sixth conception of the good is that of the “complete good.” The complete good occurs when every human being has a good will and is as happy as the normal conditions of human life allow.

What status does Kant assign his moral doctrine? According to Rawls, Kant’s doctrine of morality aspires to represent “the very existence and constitution of the order of values” (PL, 99). Kant thinks the objectivity and validity of the metaphysics of morals can and should be recognized by all reasonable and rational beings (though only from a practical point of view) (PL, 115). Kant’s notion of objectivity complements his aim to construct a moral doctrine that is both general (applying to a wide range of subjects) and comprehensive (providing a precisely articulated system of the values and virtues of human life) (PL, 13). Kant denies a reasonable person can disagree with him once the latter has reflected upon his system. Kant is “not troubled by the diversity and conflicts between our moral judgments; he supposes that what he calls ‘common human reason,’ which we all share, judges in more or less the same way” (LH, 15). At his most dogmatic, for Rawls, Kant aims to construct a secular comprehensive moral doctrine that all people can and should view as authoritative.

In *Political Liberalism*, more so than in his earlier works, Rawls emphasizes the difference between his view of the structure and content of a constructivist procedure and Kant’s. Yet Rawls’s late disavowal of Kant’s moral doctrine (at least as the basis of a stable and just society) should not divert us from noticing how closely political constructivism tracks and modifies Kant’s moral constructivism.<sup>17</sup>

The leading idea of political constructivism, as before, is to draw a connection between a conception of the person, a procedure of construction, and the principles of justice (PL, 93). The relevant conception of the person for political constructivism is that of “the citizen.” This conception is implicit in the tradition of democratic thought; it can be found by anyone who reflects upon our “everyday [political] conception of persons as the basic units of thought, deliberation, and responsibility” (PL, 18). It is important for Rawls that a political conception of the person not be drawn from any particular metaphysical doctrine (PL, 29). Kant, for example, might have based his moral doctrine on a conception of the *vernünftige Wesen überhaupt* so that the doctrine applies to God and angels as well as human beings (TJ, 226); political liberalism seeks to

---

17. Officially, political liberalism draws upon ideas circulating in the public political culture. Rawls maintains that it is possible to comprehend political liberalism without reading Kant. The purpose of political liberalism, however, is to organize familiar ideas “in a somewhat different way than before” (PL, 9), and Rawls draws heavily upon Kant to accomplish this. Kantian constructivism makes political constructivism possible.

avoid such controversial claims, at least on questions of constitutional essentials.<sup>18</sup> According to Rawls, the idea of the citizen can be found in a natural way, by thinking about how we use the idea in “everyday speech” (PL, 48). Citizens, we say, are persons who are both reasonable and rational. The term “rational” represents the capacity a citizen has for a conception of the good (PL, 19, 50–51). We assert that a person is rational when he or she is capable of identifying his or her particular ends and interests and the best means to attain or advance them. Rational is not a synonym for egoistic; rational agents may have affections for persons and attachments to communities. Yet a merely rational person does not necessarily observe the fair terms of cooperation that a just society requires. Rather, the term “reasonable” represents the capacity a citizen has for a sense of justice (PL, 19, 48–50). A reasonable person, we think, willingly proposes and honors the fair terms of cooperation and acknowledges the burdens of judgment and accepts their consequences. “Reasonable persons, we say, are not moved by the general good as such but desire for its own sake a social world in which they, as free and equal, can cooperate with others on terms all can accept” (PL, 50).

Rawls, like Kant, constructs the content of a political conception of justice by laying out a certain procedure of reasoning. The idea, again, is to uncover the concepts and principles implicit in our (political) conceptualization of the person as rational and reasonable. For both Rawls and Kant, the aim is “to formulate a procedural representation in which, so far as possible, all the relevant criteria of correct reasoning . . . are incorporated and open to view. Judgments are reasonable and sound if they result in following the correct procedure correctly and rely only on true premises” (PL, 102). The major difference between the constructivisms of Rawls and Kant concerns their characterization of “pure” practical reason. Kant often speaks as if pure practical reason were a faculty that transcends the natural world and historical contingency. Kant maintains that pure practical reason provides each reasonable and rational being with identical concepts and principles in all circumstances. Early Rawls thinks that we can characterize the reasonable better in “a theory of the moral sentiments” than in the doctrine of transcendental idealism (TJ, 44). Late Rawls holds the more chastened position that there are many acceptable ways to characterize the reasonable and that an appeal to everyday speech is simply the best way for public reason to proceed. In either case, Rawls aims to construct a non-metaphysical conception of justice (and the reasonable) that all citizens can share.

---

18. In his last works, Rawls formulated a “proviso” that allows points drawn from comprehensive moral doctrines to be introduced into public debates about the basic structure of society if they can also be subsequently reformulated in terms that accord with the idea of public reason. See IP.

One way to construct the content of a political conception of justice, according to Rawls, is to follow the sequence of the six conceptions of the good in justice as fairness (JF, 141–42; PL, 173–211).<sup>19</sup> By doing this, we gain insight into how citizens of a democratic constitutional regime may pursue their conceptions of the good within the constraints of the right (JF, 141–42). In other words, justice as fairness models the conviction of citizens in a well-ordered democratic society that the rational, though essential to our lives, must be framed by and subordinated to the reasonable. The first conception of the good is that of “goodness as rationality.” This conception accounts for our conviction that all citizens have an intuitive plan of life and an idea of how best to achieve it. Any political conception of justice, according to Rawls, must assume that human existence and the fulfillment of basic human needs and purposes are good. The second conception of the good is that of “primary goods.” Primary goods identify what all citizens need (rather than desire or prefer) in order to be free and equal persons living a complete life. Primary goods are the things that all citizens require in order to develop their moral powers and pursue their determinate conceptions of the good: the basic rights and liberties, freedom of movement, the powers and prerogatives of governance, income and wealth, and the social bases of self-respect. The idea of primary goods, though not a moral concept as such, helps us think about questions of civic and economic equality. The third conception is that of “permissible conceptions of the good.” In order to determine whether our conception of the good is permissible, we need to test it by the principles of justice as fairness. The way to perform this test is to use the model conception of the original position. The conception of the original position, in brief, asks us to imagine what principles of justice for the basic structure of society would be chosen by rational persons behind a veil of ignorance, that is by persons who do not know what place they will have in the adjusted social world the enactment of their principles helps create. The device of the original position, according to Rawls, produces two principles: one positing equality in the assignment of basic rights and duties, the other stating that social and economic inequalities are permitted only if offices and positions are open to all under

---

19. Rawls sometimes draws a sharp distinction between what parts of political constructivism are laid out and what parts are constructed. He states: in a “literal sense, only the substantive principles specifying content of political right and justice are constructed. The procedure itself is simply laid out using as starting points the basic conceptions of society and person, the principles of practical reason, and the public role of a political conception of justice” (PL, 104). How sharp can this distinction be, however, given Rawls’s understanding of the nature of moral theory? It seems that in the “back and forth” process of searching for reflective equilibrium, certain ideas will at one point be taken as the material from which we begin and at other points will be the material that we construct. For example, a substantial amount of philosophical reflection and construction of concepts is necessary to arrive at the idea of a “well-ordered society” (“a fair system of cooperation between reasonable and rational citizens regarded as free and equal”) from which, according to Rawls, political liberalism starts (PL, 103).

conditions of fair equality of opportunity and only if social and economic inequalities are to the benefit of the least-advantaged members of society (JF, 42–43). The idea of permissible conceptions of the good helps us conceptualize what ways of life a democracy ought to permit and, if possible, encourage. The fourth idea of the good in justice as fairness is that of the “political virtues.” The political virtues specify the ideal of a good citizen of a democratic state. The political virtues identify the ways that a good citizen, who also has a determinate conception of the good, embodies such characteristics as civility, tolerance, reasonableness, and a sense of fairness. The fifth and sixth conceptions of the good identify the social world that citizens ought to try to realize. The fifth conception is that of a “well-ordered society.” This is a social world where citizens respect, and act from, the principles of justice as fairness. Finally, the sixth conception is that of “a social union of social unions.” This is a world in which each citizen is just and happy, as each person benefits from the individual and communal excellences fostered by a just and stable society. Lest the analogy with Kant go unnoticed, Rawls points out that his construction of a sequence of ideas of the good “has a parallel in the way Kant can be seen to construct six ideas of the good in his moral view” (PL, 207).

What status does Rawls assign his political conception of justice? Like Kant, Rawls seeks to call his conception of justice “objective” (PL 110–29). Unlike Kant, Rawls seeks to characterize the objective without recourse to an independent order of values. Rawls identifies several features of objectivity that any political conception of justice must have. An objective political conception must establish a public framework of thought sufficient to aid our judgment. It must specify a concept of a correct judgment (either true or reasonable) made from its point of view. It must specify an order of reasons that override other reasons that citizens may have for performing an action. It must distinguish the objective from the subjective point of view and assign priority to the former. It must be able to explain how rational and reasonable people can come to an agreement in judgment. Finally, it must also be able to explain how rational and reasonable people who share an objective conception of justice may reasonably disagree. Rawls thinks that justice as fairness meets all of these criteria. Why, one might ask, should the Kantian, rational intuitionist, or citizen of faith accept this conception of objectivity? The response, according to Rawls, must focus on the limited scope of political constructivism. Political constructivism is not an account of moral values generally, but only of the values that characterize the domain of the political. To say that a political conception is objective is only to say that it is sufficient to convince other citizens that it is reasonable. A political conception of justice does not offer the “full story” about its validity. Different reasonable comprehensive doctrines, Rawls trusts, will offer different explanations for why justice as fairness should be considered objective.

## Authenticating Principles

Kant's legacy to the tradition of moral philosophy, according to Rawls, includes an account of authentication (TK, 532). This remark, from Rawls's 1989 essay on Kant, seems paradoxical. Kant addresses the question, "What kind of authentication has the moral law?" Rawls thinks this question is inappropriate for political liberalism. Rawls opts to rework Kant's account of authentication, however, rather than preserve it in its original form.<sup>20</sup>

Kant's attempts to justify (or "authenticate") the moral law, according to Rawls, can be divided into two phases (TK, 517–23). Before the *Critique of Practical Reason*, Kant sought to prove the moral law's objective and universal validity through a deduction. This deduction took several forms, each one problematic from a Kantian perspective. First, in his notebooks from the 1770s, Kant sought to ground the moral law in theoretical reason. The difficulty with this strategy, however, is that practical reason—that is, our capacity to produce objects in accordance with a conception of those objects—differs from theoretical reason—that is, our capacity to know objects given to us in sensible experience. Kant concludes that it is impossible to derive insights into the reasonable from reflecting on the rational. Second, in the *Critique of Pure Reason*, Kant entertains the possibility of providing a transcendental deduction for the moral law resembling the one given for the categories. That is, Kant considers the argument that the moral law is presupposed in our moral consciousness in the same way the categories are presupposed in our sensible experience of objects. This argument, however, collapses moral consciousness into sensible experience, in which case several conceptions of the moral law could fit our moral experience. Kant determines that we need to maintain a sharp distinction between moral consciousness—which can produce objects the world has never seen—and sensible experience—which cannot. Third, Kant, in parts of the *Critique of Pure Reason* and the *Critique of Practical Reason*, seems to outline another transcendental argument for the moral law, this time positing that just as the categories are necessarily presupposed in our unified public experience of objects, so too the moral law must be presupposed if we are to have a unified public order of conduct for a plurality of persons. The difficulty, once again, is that it is possible to conceptualize a unified public order of conduct in diverse ways. Kant has not yet proven the superiority of his moral doctrine to that of

---

20. Recent attempts to defend Kant's doctrine of the fact of reason include Henry Allison, *Kant's Theory of Freedom* (New York: Cambridge University Press, 1990), chapter 13, and Dieter Henrich, *The Unity of Reason: Essays on Kant's Philosophy*, ed. Richard L. Velkley, trans. Manfred Kuehn (Cambridge: Harvard University Press, 1994), chapter 2. Rawls draws upon Henrich to explain Kant's ideas of authentication, but Rawls does not share Henrich's apparent belief in the universal applicability of Kant's "concept of moral insight."

utilitarianism, perfectionism, or intuitionism. Finally, in Section III of the *Groundwork*, Kant tries to derive the moral law from the idea of freedom. The difficulty here is that it is impossible to have an intellectual intuition of freedom. All of these arguments, and their respective flaws, lead Kant to surmise that the moral law cannot be given a deduction.

In the *Critique of Practical Reason*, Kant expounds another strategy of authentication that holds more promise for his purpose and ours (according to Rawls). The basis of a “coherentist account” of authentication is to show the place and role of each form of reason in the constitution of reason as a whole. Within this overarching framework, each form of reason has a different authentication. The categories of the understanding are authenticated because they make possible our experience of objects in space and time. Pure speculative reason is authenticated by its role in regulating the understanding and organizing our empirical knowledge into a unity. Empirical practical reason is authenticated by its role in regulating our desires and inclinations into a conception of happiness. Pure practical reason, finally, is authenticated by the “fact of reason.” This is the fact that in our common moral consciousness we recognize and affirm that the moral law is “supremely authoritative and immediately directive” for us. Pure practical reason also gains a “credential” by its role as the *ratio cognoscendi* of freedom. The doctrine of the fact of reason has often been criticized as a regress into intuitionism or dogmatism. Rawls’s position, however, is that a “coherentist doctrine of practical reason is not without strengths as a possible view” (TK, 523).

Rawls abjures the idea, just as late Kant did, that practical philosophy can justify a theory of justice through a theoretical deduction. “I do not claim for the principles of justice proposed that they are necessary truths or derivable from such truths. A conception of justice cannot be deduced from self-evident premises or conditions on principles” (TJ, 19). Rawls also rejects the idea that a theory of justice should be grounded on a moral fact able to command the assent of all reasonable and rational citizens. That would be a step into dogmatism, especially given the fact of reasonable pluralism in modern democratic societies. Rawls, instead, articulates an idea of public justification that can be shared by all reasonable comprehensive doctrines (JF, 26–29). What can this idea draw from Kant?

Rawls terms his naturalized Kantian account of authentication the idea of reflective equilibrium.<sup>21</sup> Rawls defends his theory of justice because it “best fits all our considered convictions on reflection and organizes them into a coherent

---

21. I take the phrase, “naturalizing Kant,” from Paul Guyer, “Naturalizing Kant,” in *Kant verstehen: Understanding Kant*, ed. Dieter Schönecker and Thomas Zwenger (Darmstadt: Wissenschaftliche Buchgesellschaft, 2001), 59–84.

view" (JF 31; see also TJ, 18). The key philosophical difference between Rawls and Kant, once again, is how to conceive of the dualism between our sensible and our *vernünftige* (reasonable/rational) natures. Kant's view is that this dualism is deep, and that moral philosophy should appeal strictly to reason. Kant thinks that an explanation of the unity of reason suffices to authenticate the moral law to all rational beings. Rawls, we have seen, seeks to overcome Kant's dualism between our sensible and rational natures. Kant's approach, for Rawls, lacks "explicable connections with human conduct" (TJ, 226). Rawls believes that our sensible and rational natures intertwine and that the idea of public justification should be conceived of accordingly.

Rawls describes the process of reflective equilibrium as follows (JF, 29–32). We begin by gathering together "considered political judgments at all levels of generality," including judgments about particular actions, individuals, institutions, and social policies. Although these political judgments are considered, in the sense that they are at least proto-rational, they issue, so to speak, from the heart. An example of such a judgment is Abraham Lincoln's belief that "if slavery is not wrong, nothing is wrong." Next, we find a conception of political justice that makes the fewest revisions in our initial political judgments. At this point, we are in "narrow" reflective equilibrium. Our (primarily sensible) judgments and our (primarily rational) concepts and principles of justice cohere between and among themselves. The problem with narrow reflective equilibrium is that it is premised upon ignorance of available theoretical options. The subsequent step in the process, therefore, is to seek "wide" reflective equilibrium. This occurs when we examine alternate ideas of justice and possibly revise our preferred political conception of justice. When we attain or approach a state of wide reflective equilibrium, we believe that our political conception of justice is the most reasonable one for us. This belief is supported if we trust other citizens could share this conception and thus help society attain "general" or "full" reflective equilibrium. According to Rawls, justice as fairness best characterizes our position in wide and general reflective equilibrium. Justice as fairness, alone among contemporary alternatives, "meets the need for a basis of public justification on questions of political justice" (JF 32).

Rawls proposes the idea of reflective equilibrium to "you and me" deciding whether justice as fairness is the best available account of justice (PL, 28). Rawls introduces two other ideas to think about how a political conception of justice could be justified to citizens of a well-ordered society. One of these (the idea of public reason) has a Kantian pedigree; the other (the idea of an overlapping consensus) does not.

The idea of public reason, as noted above, identifies the way that citizens of a well-ordered society can justify their political conception of justice to one another. This idea reconstructs an idea formulated by Kant in the essay "What is

Enlightenment?" For Kant, public reason is "the use which someone makes of it as a scholar before the entire public world of readers."<sup>22</sup> Kant uses the idea of public reason, presumably, as part of his overall project of justifying Enlightenment liberalism to all citizens of the world. For Rawls, on the other hand, public reason identifies "the kinds of reasons [citizens] may reasonably give one another when fundamental political questions are at stake" (IP, 574). In an inversion of Kant's project, Rawls seeks to take comprehensive doctrines of truth or right out of the public sphere.<sup>23</sup> For Rawls, the idea of public reason places limits on how citizens—most importantly, government officials—can publicly justify a political conception of justice to one another. Three criteria, according to Rawls, constrain public reason (IP, 575). First, the target of public reason may only be the reason of free and equal citizens (as opposed to their prejudices). Second, the subject of public reason may only be constitutional essentials and matters of basic justice. Third, the nature and content of public reason must themselves be public. All citizens must, if they wish, be able to discover how the political conception of justice is justified. What makes a liberal political conception of justice better able than alternatives to satisfy the criteria of public reason? It alone can be "worked out from fundamental ideas seen as implicit in the public political culture of a constitutional regime" (IP, 584). Only a liberal conception of justice—of which justice as fairness is an example—can be publicly justified without recourse to comprehensive moral doctrines.

Rawls introduces the idea of an overlapping consensus in order to help us think further about how citizens in a well-ordered society could conceive of the idea of public justification. The idea of an overlapping consensus helps "make the idea of a well-ordered society more realistic and to adjust it to the historical and social conditions of democratic societies" (JF, 32). Kant's hope, in his time, was to achieve a consensus of reasonable and rational beings on the objectivity and validity of his comprehensive moral doctrine. Several facts make this and similar hopes unreasonable for us (JF, 33–36). The most important is the fact of reasonable pluralism. This is the fact that the free exercise of human reason in a free democratic society leads to a diversity of conflicting and irreconcilable, though reasonable, comprehensive doctrines. Another is the fact of oppression. This is the fact that in any modern society a consensus on any one comprehensive moral doctrine can only be achieved through the use of state oppression. Next, there is the fact that any enduring and secure democratic regime must be supported by a majority of its citizens.

---

22. Immanuel Kant, "What is Enlightenment?" (1787), 8:37, 18.

23. For a discussion of how Kant's idea of public reason inverts the traditional liberal notion of freedom of conscience, see Michel Foucault's essay "What is Enlightenment?" in *The Foucault Reader*, ed. Paul Rabinow (New York: Pantheon Publishers, 1984), 32–50.

This support could not be guaranteed if the officially sanctioned conception of justice were based on a sectarian moral doctrine. Finally, there is the fact that many people are beginning to see the real possibility of a society that is both just and pluralistic. All of these facts emerge from and reinforce our awareness that the burdens of judgment preclude agreement among reasonable and rational people on a single comprehensive moral doctrine. What is Rawls's hope, then, in introducing the idea of an overlapping consensus? This idea helps reconcile us to the fact that a democratic society cannot be a community as Kant and others once envisioned (JF, 3). More positively, this idea is "realistically utopian" (JF, 4). The idea of an overlapping consensus helps us imagine how a well-ordered society faced with the fact of reasonable pluralism may be unified and stable. Our dreams, for Rawls, differ from Kant's.

### **Rawls and the Kantian Ethos**

Rawls, we have seen, distances his theory of justice from Kant's moral doctrine and Enlightenment liberalism. Yet Rawls takes innumerable ideas and strategies from Kant and the Enlightenment and recasts them for his own purposes. Rawls's critics—such as Allan Bloom, Michael Sandel, and Bonnie Honig—and Kant scholars—such as Larry Krasnoff, Onora O'Neill, and Allen W. Wood—sometimes accuse Rawls of misrepresenting Kant when he originates concepts such as the "CI-Procedure" or "Kantian constructivism." Rawls maintains, however, that he is Kantian when he refashions Kant's ideas or creates new ones. For Rawls, the Kantian ethos (or spirit) impels us to exercise in our time the philosophical courage that Kant exercised in the eighteenth century.

Rawls explains his approach to reading Hume, Kant, and other Enlightenment philosophers in his *Lectures on the History of Moral Philosophy*:

One of the benefits of studying historical texts . . . is that we come to see how philosophical questions can take on a different cast from, and are indeed shaped by, the scheme of thought from within which they are asked. And this is illuminating, not only in itself, as it discloses to us different forms of philosophical thought, but also because it prompts us to consider by contrast our own scheme of thought, perhaps still implicit and not articulated, from within which we now ask our questions. And this self-clarification helps us to decide which questions we really want to resolve, which ones we can reasonably expect to settle, and much else (LH, 17–18).

Rawls respects traditional philosophical exegesis. He values scholarship that clarifies the ideas and systems of profound and difficult philosophers. Rawls also

echoes the Enlightenment's disdain for scholasticism, or spurious erudition.<sup>24</sup> Kant addressed different questions than the ones that interest Rawls. Rawls appreciates the mental exercise of reading obscure passages in Kant's corpus, but he does not want to lose sight of the problems of the twenty-first century. For Rawls, the real value of studying Kant's scheme of thought is to help clarify our own thinking. In the Preface to the *Groundwork*, for example, Kant states that the problem of practical philosophy is to search for and establish the supreme principle of morality. Rawls realizes that this is not the problem of late-modern pluralistic societies. Yet Rawls takes one of Kant's key concepts—autonomy—to formulate the problem of political liberalism. Rawls learns from Kant, but he is not beholden to him.

In the reminiscence that serves as the forward to the *Lectures on the History of Moral Philosophy*, Rawls explains how Kantians ought to extend Kant's legacy. In the *Critique of Pure Reason* (B866), Kant states:

'[W]e cannot learn philosophy; for where is it, who is in possession of it, and how shall we recognize it? We can only learn to philosophize, that is, to exercise the talent of reason . . . on certain actually existing attempts at philosophy, always, however, reserving the right of reason to investigate, to confirm, or to reject these principles in their very sources.' Thus we learn moral and political philosophy . . . by studying the exemplars, those noted figures who have made cherished attempts at philosophy; and if we are lucky, we find a way to go beyond them (LH, xvii).

For Rawls, we study great philosophers to train our minds to raise important questions, make sophisticated arguments, link ideas, and so forth. Rawls cherishes Kant's *Groundwork* and Hume's *Treatise*. Participants in the Enlightenment tradition, however, do not treat any earlier doctrine as dogma. For Kant, philosophers always possess the right to weigh principles. That is, philosophers are entitled to discard old ideas and create new ones. Rawls embodies the Kantian ethos when he goes beyond Kant.<sup>25</sup>

In *Political Liberalism*, Rawls no longer calls his theory of justice Kantian. He states that he is more attuned to the fact of pluralism than he was in *A Theory of Justice* (PL, xix). This explanation does not seem to suffice. Early Rawls

24. See Peter Gay, *The Enlightenment: An Interpretation. Volume One: The Rise of Modern Paganism* (New York: Norton Library, 1966), 226–36.

25. Rawls also expresses this idea in his 1975 essay, "The Independence of Moral Theory":

All the main conceptions in the tradition of moral philosophy must be continually renewed: we must try to strengthen their formulation by noting the criticisms that are exchanged and by incorporating in each the advances of the others, so far as this is possible. In this endeavor the aim of those most attracted to a particular view should be not to confute but to perfect (IM, 302).

constructs a theory that may be endorsed by a diversity of social unions (TJ, 462). Early Rawls draws upon Kant but he does not try to convert all citizens to Kant's idea of the good. Late Rawls does not seem appreciably less Kantian than early Rawls. Rawls's late essay on Kant and his lectures on Kant attest that nearly every Rawlsian idea has its roots in Kantianism. Rawls may have rejected the Kantian moniker for another reason: he became impatient with proving his Kantian credentials. Rawls's critics, including Bloom, Sandel, Honig, and orthodox Kantians repeatedly point out "mistakes" in Rawls's Kant scholarship. Rawls, I conjecture, did not want political liberalism to succeed or fail based on scholarly criteria. Yet one has only to read Rawls's late reminiscence on teaching Kant to see that Rawls drew inspiration from Kant his entire life. Rawls lived the Kantian ethos even as he repudiated Kant's moral doctrine as a political conception of justice.

What is the best way to continue the Enlightenment today? Many Kant scholars argue that it is to clarify and defend Kant's doctrines. Rawls would applaud the recent boom in Kant scholarship, but he would also warn against adulating Kant (LH, xviii). For Rawls, the Kantian ethos encourages us to think for ourselves, to create ideas to address the problems of our time. For Rawls, heirs of the Enlightenment ought to rewrite Kant's *Groundwork* and, potentially, *Justice as Fairness*.