

The London School of Economics and Political Science

Political Theory and Social Practices

G.A. Cohen, Rawls, Habermas and the problem
of self-grounding

James Gledhill

A thesis submitted to the Department of Government of the
London School of Economics for the degree of Doctor of
Philosophy, London, July 2010

Declaration

I certify that the thesis I have presented for examination for the MPhil/PhD degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

The copyright of this thesis rests with the author. Quotation from it is permitted, provided that full acknowledgement is made. This thesis may not be reproduced without the prior written consent of the author.

I warrant that this authorization does not, to the best of my belief, infringe the rights of any third party.

Abstract

In a time of transitions, post-Rawlsian political philosophy is itself in transition, engaged in a methodological dispute regarding the relationship between political theory and changing social practices. This thesis enters this dispute through engaging with John Rawls's philosophical project and the two leading but contrasting critiques of Rawls's constructivist methodology. I first seek to rescue constructivism from G.A. Cohen's critique of its fact-dependence, but secondly argue with Jürgen Habermas for a shift from constructivism to reconstructivism. Part I establishes a theoretical framework. I analyse competing paradigms of the relationship between normative principles and social practices and situate them in relation to the problem of self-grounding. This is the methodological problem of how, in accordance with a conception of freedom as autonomy, philosophy can find normative foundations within existing social practices. While Cohen rejects this problem in arguing for a choice between realism and utopianism, Rawls's realistic utopianism and Habermas's utopian realism are both driven by the problem of self-grounding. Part II defends Rawls's constructivism against Cohen's criticism of its restricted focus on the basic structure of society and fact-dependence. Part III analyses and critiques the development of Rawls's project. It analyses Rawls's concern with the problem of stability and critiques from a Habermasian perspective the approach to the problem of self-grounding this represents. Part IV argues that post-Rawlsian deliberative democrats who have sought to combine ideas from Rawls and Habermas also fail to adequately address this methodological problem. Part V engages with Habermas on his own terms. I first analyse Habermas's reconstruction of the tension between facticity and validity in morality and politics. On this basis, I conclude that Habermas's procedural reconstructivism allows him to more successfully address the problem of self-grounding than Rawls's substantive constructivism, and assess the implications of this conclusion in theory and practice.

Acknowledgements

My thanks to my doctoral supervisor Katrin Flikschuh, both for her encouragement and for a probing scepticism that challenged me to articulate the ideas in this thesis with greater clarity, and to my advisor Paul Kelly for his guidance. I am grateful to my undergraduate tutors Elizabeth Frazer and Stephen Mulhall for setting me on this road and for their help since.

I have presented chapters of this thesis at a number of conferences and seminars and I benefited from valuable feedback and stimulating discussions. Particular thanks to Catherine Audard, Gordon Finlayson, Jürgen Habermas, Robert Jubb, Pietro Maffettone, Sebastian Stein and Federico Zuolo.

I could not have asked for a better group of fellow doctoral students to work alongside. Thanks to Muriel Kahane, Camillia Kong and Jeremy Williams for their camaraderie and intellectual inspiration.

My wife Alejandra has been there every step of the way offering more love, support and encouragement than I could have ever hoped for. She will know that these few words express a lot of emotions.

This thesis is dedicated to my parents. Their love and unconditional support has been the constant background to this as with everything I have done.

Contents

Abbreviations	8
Introduction: Political Theory in a Time of Transitions	11
PART I. IDEAL THEORY, CRITICAL THEORY AND THE PROBLEM OF SELF-GROUNDING	15
1. The Idealism of Ideal Theory	16
1.1 The Fact-sensitivity Spectrum	17
1.2 Realism vs. Moralism	19
1.3 Realistic Utopianism	21
1.4 The Pursuit of the Ideal within our Social World	29
1.5 From Realistic Utopianism to Utopian Realism	35
2. Post-Kantian Constructivism and the Problem of Self-grounding	39
2.1 Rousseau and the Counterfactual Imagination	40
2.2 Kant and Self-legislation	43
2.3 Pragmatism and Immanent Foundationalism	45
2.4 Hegel and the Philosophical Problem of Modernity	51
2.5 Rawls's Hegelianized Kantianism	54
2.6 Marx and Socialism, Utopian and Scientific	58
3. Bringing Rawls's and Habermas's Self-grounding Projects into Dialogue	63
3.1 The Rawls-Habermas Debate	63
3.2 Kantian Family Dispute or Self-grounding Family Resemblance?	65
3.3 Rawls: Political Liberalism as Hegelianized Kantianism	68
3.4 Habermas: Kantian Republicanism as Kantianized Hegelianism	69
PART II. CONSTRUCTIVISM BEYOND REALISM AND UTOPIANISM	72
4. The Basic Structure and the Division of Justificatory Labour	73
4.1 Platonic and Aristotelian Ways to Think About Justice	75
4.2 Justice of the Basic Structure and Justice of Personal Behaviour	78
4.3 From the Division of Moral Labour to the Division of Justificatory Labour	82
4.4 Autonomously Realizing Autonomy	86
4.5 Justificatory Community and Individual Self-realization	89
5. Rescuing Constructivism from Platonism and Positivism	95
5.1 Platonic Atomism and Quinean Holism	96
5.2 Facts and Principles	100
5.3 Beyond Facts and Principles	102
5.4 Principles of Justice, Rules of Regulation and Principled Rules of Regulative Justice	109
5.5 Constructivism as Realistic Utopianism	111
PART III. RAWLS'S REALISTIC UTOPIANISM	119
6. The Problem of Self-grounding in the Search for Stability	120
6.1 A Self-grounding Constructivist Blueprint for Reading Rawls	121

6.2 The Elements of Justice as Fairness	124
6.3 A Theory of Justice <i>In Mediis Rebus</i>	130
6.4 Ideal Feasibility and Utopianism	137
6.5 Real Feasibility and the Search for Stability for the Right Reasons	141
6.6 Realizability and Reasonable Pluralism	143
7. Overlapping Consensus and the Unrealizability of Reconciliation	147
7.1 Overlapping Consensus and Self-grounding Justification	148
7.2 Reconciliation by Public Reason	152
7.3 The <i>Deus ex Machina</i>	158
7.4 Post-Rawlsianism	164
7.5 On the Very Idea of a Comprehensive Doctrine	167
PART IV. BEYOND ‘RAWLBERMAS’	171
8. The Democracy/Contractualism Analogy Beyond ‘Rawlbermas’	172
8.1 The Democracy/Contractualism Analogy	174
8.2 Contractarianisms and Contractualisms	176
8.3 Mirroring a Deliberative Initial Situation	180
8.4 Beyond ‘Rawlbermas’	184
8.5 Rescuing Proceduralism	187
PART V. HABERMAS’S UTOPIAN REALISM	194
9. Reconstructivism and Transcendence from Within	195
9.1 Beyond Rawls and Luhmann	196
9.2 Communicative Reason and Reflexive Social Practices	200
9.3 Kantianized Hegelianism and Transcendence from Within	207
9.4 Truth, Rightness and Truthfulness	210
9.5 A Discourse Theory of Law and Democracy	215
10. From Procedure in Substance to Substance in Procedure	224
10.1 Procedure and Substance in Justice and in Practical Reason	225
10.2 Rawls’s Constructivism: Procedure in Substance	230
10.3 Habermas’s Reconstructivism: Substance in Procedure	233
10.4 Procedural Rationality and Religious Pluralism	237
10.5 The Project of Self-grounding and the Exhaustion of Utopian Energies	239
Conclusion: Philosophical Proceduralism and Political Radicalism in a Time of Transitions	244
References	248

List of Figures

Figure 1. The Fact-sensitivity Spectrum	18
Figure 2. Platonic, Kantian and Rousseauian Paradigms of the Relationship between Theory and Practice	22
Figure 3. Self-grounding Projects: Rawls's Hegelianized Kantianism and Habermas's Kantianized Hegelianism	68
Figure 4. Procedural and Substantive Justifications of Democracy	175
Figure 5. Models of the Relationship between Democratic Deliberation and Justice	190

List of Tables

Table 1. Reason, Facts and Realization in the Platonic, Kantian and Rousseauian Paradigms	38
Table 2. The Rawls-Habermas Debate	147
Table 3. Linguistically Mediated Interaction	206
Table 4. Forms of Procedural Justice	231

Abbreviations

G.A. Cohen

HLF *History, Labour, and Freedom*. Oxford: Oxford University Press, 1988.

IYE *If You're an Egalitarian, How Come You're So Rich?* Cambridge, MA: Harvard University Press, 2000.

KMTH *Karl Marx's Theory of History: A Defence*. Princeton: Princeton University Press, 2001. Expanded ed. (Originally published 1978).

RJE *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press, 2008.

SFE *Self-Ownership, Freedom, and Equality*. Cambridge: Cambridge University Press, 1995.

WNS *Why Not Socialism?* Princeton: Princeton University Press, 2009.

John Rawls

CP *Collected Papers*. Cambridge, MA: Harvard University Press, 1999.

JF *Justice as Fairness: A Restatement*. Cambridge, MA: Harvard University Press, 2001.

LHMP *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press, 2000.

LHPP *Lectures on the History of Political Philosophy*. Cambridge, MA: Harvard University Press, 2007.

LP *The Law of Peoples*. Cambridge, MA: Harvard University Press, 1999.

MSF *A Brief Inquiry into the Meaning of Sin and Faith*. Cambridge, MA: Harvard University Press, 2009.

PL *Political Liberalism*. 2nd ed. New York: Columbia University Press, 1996 (includes Rawls's 'Reply to Habermas' as Lecture IX).

SGEK A Study in the Grounds of Ethical Knowledge: Considered with Reference to Judgments on the Moral Worth of Character. Ph.D. Dissertation, Princeton University, 1950.

TJ *A Theory of Justice*. Cambridge, MA: Harvard University Press, original ed. 1971, rev. ed. 1999. References are given in the form (x/y) where x is the original edition and y is the revised edition.

Jürgen Habermas

AS *Autonomy and Solidarity*. London: Verso, 1992.

BFN *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Cambridge: Polity Press, 1996.

BNR *Between Naturalism and Religion*. Cambridge: Polity Press, 2008.

BR *A Berlin Republic: Writings on Germany*. Cambridge: Polity Press, 1997.

CES *Communication and the Evolution of Society*. Cambridge: Polity Press, 1979.

DW *The Divided West*. Cambridge: Polity Press, 2006.

EFP *Europe: The Faltering Project*. Cambridge: Polity Press, 2009.

FHN *The Future of Human Nature*. Cambridge: Polity Press, 2003.

HR 1 Reconciliation through the Public Use of Reason. In *The Inclusion of the Other*. Cambridge: Polity Press, 1998.

HR 2 ‘Reasonable’ versus ‘True,’ or the Morality of Worldviews. In *The Inclusion of the Other*. Cambridge: Polity Press, 1998.

IO *The Inclusion of the Other*. Cambridge: Polity Press, 1998.

JA *Justification and Application: Remarks on Discourse Ethics*. Cambridge: Polity Press, 1994.

KHI *Knowledge and Human Interests*. Cambridge: Polity Press, 1987.

LC *Legitimation Crisis*. Cambridge: Polity Press, 1975.

LM Law and Morality. *The Tanner Lectures on Human Values*. 1986.

LSS *On the Logic of the Social Sciences*. Cambridge: Polity Press, 1988.

MCCA *Moral Consciousness and Communicative Action*. Cambridge: Polity Press, 1990.

NC *The New Conservatism*. Cambridge: Polity Press, 1990.

OPC *On the Pragmatics of Communication*. Cambridge: Polity Press, 1998.

OPSI *On the Pragmatics of Social Interaction*. Cambridge: Polity Press, 2001.

PC *The Postnational Constellation*. Cambridge: Polity Press, 2001.

PDM *The Philosophical Discourse of Modernity*. Cambridge: Polity Press, 1987.

PF *The Past as Future*. Cambridge: Polity Press, 1994.

PP *Philosophical-Political Profiles*. Cambridge, MA: MIT Press, 1983.

PT *Postmetaphysical Thinking*. Cambridge, MA: MIT Press, 1992.

RC A Reply to my Critics. In *Habermas: Critical Debates*, edited by J. B. Thompson, and D. Held. Cambridge, MA: MIT Press, 1982.

RR *Religion and Rationality: Essays on Reason, God, and Modernity*. Cambridge: Polity Press, 2002.

RSP Reply to Symposium Participants, Benjamin N. Cardozo School of Law. In *Habermas on Law and Democracy: Critical Exchanges*, edited by M. Rosenfeld. Berkeley and Los Angeles: University of California Press, 1998.

STPS *The Structural Transformation of the Public Sphere*. Cambridge: Polity Press, 1989.

TaJ *Truth and Justification*. Cambridge: Polity Press, 2003.

TCA 1 *The Theory of Communicative Action: Vol.1, Reason and the Rationalization of Society*. Cambridge: Polity Press, 1984.

TCA 2 *The Theory of Communicative Action: Vol.2, A Critique of Functionalist Reason*. Cambridge: Polity Press, 1987.

TP *Theory and Practice*. Cambridge: Polity Press, 1973.

TT *Time of Transitions*. Cambridge: Polity Press, 2006.

The history of political theory is not the history of different answers to one and the same question, but the history of a problem more or less constantly changing, whose solution was changing with it.

R.G. Collingwood, *An Autobiography*

Quoted by John Rawls in *Lectures on the History of Political Philosophy*

The rose in the cross of the present may have grown pale, but it is not yet completely faded.

Jürgen Habermas, ‘From Kant to Hegel and Back Again:
The Move toward Detranscendentalization’

Introduction: Political Theory in a Time of Transitions

‘The more things change, the more they remain the same’. So Jürgen Habermas summarises the mood of the mid-1980s, exemplified by Arnold Gehlen’s idea of *posthistoire*. History has reached a ‘crystalline’ state in which nothing really surprising can happen anymore; the Enlightenment is dead, ‘the history of ideas has concluded’ (Gehlen, quoted in PDM: 3). But writing in 1990, and in marked contrast to Francis Fukuyama’s (1989, 1992) continued proclamation of the end of history, Habermas could write:

History is once again on the move, accelerating, even overheating. New problems are shifting the old perspectives. What is more important, new perspectives are opening up for the future, points of view that restore our ability to perceive alternative courses of action (BFN: 491).

Elaborating on this theme, more recently Habermas has presented a *Zeitdiagnose* through the motif of a ‘time of transitions’, an analysis that encompasses the processes of German unification, European integration and globalization (TT). Expressions of a sense of transition are of course a familiar refrain when it comes to capturing the spirit of the modern age (Mill 1997 [1831]). Indeed, for Habermas these transition processes are part of a broader process of modernization, both Western social modernization and the claim to universalism of the associated philosophical discourse of modernity. The philosophical problem of modernity is, as Habermas puts it following Hegel, the problem of self-grounding. Modernity cannot appeal to an exemplary past but ‘*has to create its normativity out of itself*’ (PDM: 7).¹ In our time of transitions, philosophy can have no recourse to external ‘rails’ to arrest a vertiginous sense of contingency and incompleteness. Faced with processes of transition that are ‘stalled’ (TT: xv) or ‘faltering’ (EFP), the task for philosophy is to identify potentials in the present that maintain an openness to the horizon of the future.

In response to this time of transitions, political theory is itself in transition (O’Sullivan 2000a). This is not to discount the perception that ‘political theory is *always* in transition’ (O’Sullivan 2000b: 1), but in a manner reminiscent of Thomas Kuhn’s (1970) distinction between periods of ‘normal science’ puzzle-solving and revolutionary periods of paradigm reassessment, political philosophy may operate in two modes. In the first mode, theory development is driven by an ‘internal dialectics of moral learning’ as

¹ Unless otherwise stated, emphasis in quotations is in the original.

the implications of a theoretical position are developed, criticised and revised, while in the second, moral and political thinking adopts a more reactive attitude, responding to transformations in social conditions (Peritz 2003: 40-42). Much of contemporary political philosophy – whether postsocialist theories of justice and democracy, theories of multiculturalism, or of global and intergenerational justice – operates in this second mode. Since the seminal moment in the constitution of the ‘disciplinary matrix’ (Kuhn 1970) of contemporary political philosophy was the publication of John Rawls’s *A Theory of Justice* (*TJ*) in 1971, it is not surprising that with political theory in transition, Rawls’s methodology has become the focus of a *Methodenstreit*. As Marc Stears observes, ‘after decades of taking the purposes and the methods of the subdiscipline almost for granted . . . political theorists have been engaged in increasingly bitter disagreements’ (Stears 2005: 325-26; see also White 2004; Leopold 2008).

Political philosophy has an increasingly bifurcated focus, engaging with questions of global and transnational justice and democracy on the one hand, and questions of personal justice and small-scale deliberative democracy on the other. Orthogonal to this distinction in focus is an increasing methodological polarization between fact-independent pure theory and fact-sensitive contextualism. With political theory in transition, the shape of post-Rawlsian political philosophy is currently being determined, and we are faced with the prospect of a split between abstract philosophical reflection and practical political engagement. A changing tide that has brought a move from engagement with philosophical theory to a concern for relevance to political practice² threatens to leave the high grounds of political philosophy populated by only the most fervent advocates of ideal philosophical theory.

To move forward in our time of transitions, we need to know how we got to where we are. In the explosion of concern with global justice, for example, we are yet to properly address the question of whether the coherence given to the values of freedom, equality and self-determination by the institutions of the nation-state can be expanded, or even maintained, in a postnational system. In this case, as in others, in avoiding the twin pitfalls of utopian and realist political thought, ‘the hardest thing of all’, as David Runciman has put it, ‘is to find a way of harnessing the possibilities contained in the present’ (Runciman 2006: 173-74). In this thesis I focus on the two leading contemporary methodological approaches in political philosophy that take as their problem harnessing the possibilities contained in the present: Rawls’s realistically utopian constructivism and what I call Habermas’s utopian realist reconstructivism. I outline Rawls’s constructivism in relation to the criticisms of the theorist who is,

² See, for example, Barry (2005); Nussbaum (2000, 2006); Wolff (2007); Sen (2009); Archard (2009).

alongside Habermas, the most significant recent critic of Rawls, G.A. Cohen. Cohen's avowedly utopian approach to political philosophy forms the basis of an argument for socialism that Cohen had previously made on the basis of a scientific conception of historical materialism. I seek to rescue Rawls's constructivism from Cohen's Procrustean bed of utopianism versus realism, and I will argue that Rawls's position has greater affinities with Habermasian critical theory than it does with Cohen's brand of analytic philosophy. Drawing on Habermas, however, an immanent critique of Rawls's project provides reasons for endorsing a Habermasian paradigm shift in philosophical methodology from a substantive constructivism to a procedural reconstructivism.

Pursuing such methodological problems is crucial if we are to orientate our thinking in our time of transitions, but this is not a matter of first getting our methods right and then applying them to problems: 'In good Hegelian terms, the formation of basic concepts and the treatment of substantive issues belong inseparably together' (TCA 1: xli).

PART I. IDEAL THEORY, CRITICAL THEORY AND THE PROBLEM OF SELF-GROUNDING

1. The Idealism of Ideal Theory

John Rawls's view that political philosophy should work within ideal theory has been subject to increasing criticism.³ It is claimed that no amount of 'tinkering' with Rawls's theory of justice will resolve the fundamental deficiency of the 'principled paradigm of ideal theory' and that we need a paradigm shift from ideal to nonideal theory (Farrelly 2007a: xiii-xiv). Egalitarian theorists are 'beginning to think that greater attention to issues of "nonideal" theory and "the real world" is needed, following the thought that political philosophy must start from where we are' (Wolff 2007: 135). For on this account, Rawls's practice of theorizing about political institutions under the ideal theory assumption of strict compliance 'represents a potentially serious limitation on the relevance of political philosophy for real-world policy' (Brennan 2005: 258).⁴

There are, however, reasons to be cautious before condemning Rawls's conception of ideal theory. In particular, it would be premature to equate Rawls's ideal theory with the fact-independent pure theory defended by G.A. Cohen and overlook the methodological framework within which Rawls's substantive principles of justice are defended. We might start from Rawls's observation that a 'liberal political philosophy which, of course, accepts and defends the idea of constitutional democracy, is not to be seen as a theory' (LHPP: 1). A conception of justice for a democratic society is not the application of an external, theoretically derived ideal moral blueprint to a comprehensive range of subjects in the sphere of political practice. For Rawls, neither political philosophy nor justice as fairness is applied moral philosophy in this sense (JF: 14, 182; CP: 482).

I will argue against representing the ideal theory debate in terms of a spectrum of fact-sensitivity as it obscures fundamental questions about the relationship of theory to practice, focusing in particular on Bernard Williams's argument for realism in political theory. In contrast to the fact-sensitivity spectrum view, I propose that the ideal theory debate is best understood in terms of competing Platonic, Kantian and Rousseauian paradigms of the relationship between theory and practice. I set out Rawls's realistically utopian Rousseauian conception of political philosophy and show how in ideal theory Rawls constructs a practical and reasonably realistic conception of a well-ordered

³ I will not seek to distinguish between political theory and political philosophy, but see Griffin (1993) in relation to Rawls.

⁴ See Brighouse (2004: 11-30); Fabre (2007); McKinnon (2008). In Nielsen (2006), the need to take seriously nonideal theory and feasibility constraints is emphasised by Appiah (9-10), Buchanan (33-34), Galston (48-49), Mason (78-79), Pettit (116), Williams (191), Wolff (205); and in Petersen (2007) by Anderson (7-8) and Heyd (51-52).

society. Within this framework, nonideal theory considers whether and how a conception constructed in ideal theory can inform the judgements in practice of political practitioners. Before abandoning the paradigm of ideal theory it is important to understand the deep structure of Rawls's approach. Doing so enlarges the current agenda for political philosophy and brings Rawls's work into an engagement with Jürgen Habermas's, which I present as offering an alternative elaboration of the Rousseauian paradigm.

1.1 The Fact-sensitivity Spectrum

Ideal theory and nonideal theory are commonly pictured as lying at opposite ends of a fact-sensitivity spectrum (Mason 2004; Farrelly 2007a). At the nonideal end of the spectrum are those such as John Dunn who argue that the consideration of political ideals should be set within the context of what is feasible given our social, political and economic circumstances. By contrast, at the ideal end of the spectrum are those such as Cohen who defend the existence of fact-free principles of justice whose logical status and normative appeal are unaffected by issues of feasibility and constraints of human nature (Figure 1). Rawls is seen as occupying an intermediate position where ideal theory is subject to 'moderately strong feasibility constraints' (Mason 2004: 265). However, on the fact-sensitivity spectrum view, Rawls's 'moderate ideal theory' appears arbitrary and unstable. The intermediate terrain that Rawls seeks to occupy seems an

amphibious world, which contains just enough land of real social reference to avoid the tricky depths of first philosophy (the gesture is roughly: let's start out from where we're at . . .), while floating carefully enough on the waters of abstraction to avoid contact with the ground of actual political change (Anderson 2005: 108-9).

Stuck in an apparent no-man's land, Rawls's methodology has been assailed by a pincer movement. While Cohen criticises its fact-sensitivity, Amartya Sen (2006, 2009) sees Rawls as adopting a 'transcendental' approach to justice that in the face of real-world injustices can only propose a radical jump to a perfectly just world.

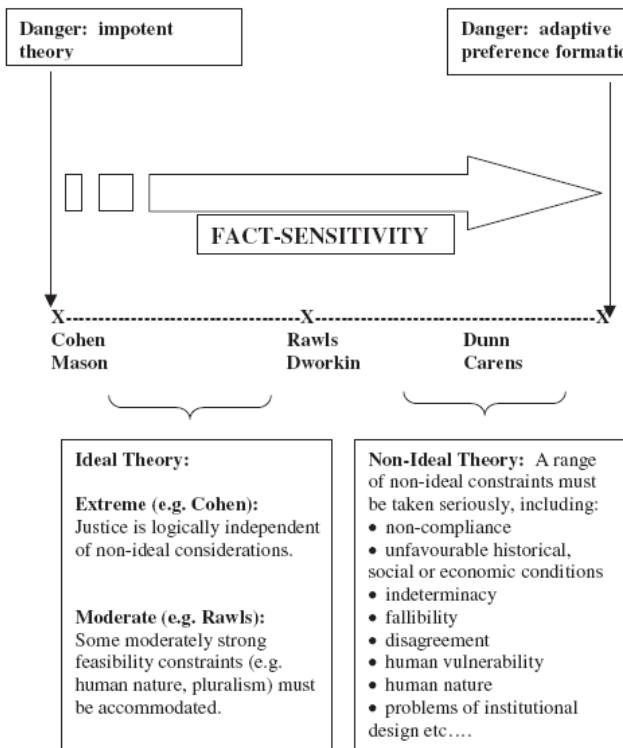


Figure 1. The Fact-sensitivity Spectrum

Source: Farrelly (2007b: 847). See also Farrelly (2007a: 51).

To see that the fact-sensitivity spectrum poses a false choice between ideal theory and nonideal theory, we need only consider the way in which presenting Dunn as an advocate of nonideal theory misconstrues the significance of his position. Dunn is not so much concerned with the fact-sensitivity demanded of political theories as with the insensitivity of the facts of the causal setting of politics to political theories, contrasting unfavourably the theoretical wisdom claimed by political theorists (professional social theorists) with the virtue of prudence possessed by political actors (amateur social theorists). This represents a fundamental realist challenge to ideal theory approaches to political philosophy that consist in ‘the application of ethics to political topics’ (Dunn 1990: 195), a challenge that has been made polemically by Raymond Geuss (2008), and that is also associated with Isaiah Berlin (2003) and John Gray (1994). It is a challenge to which views located within a political-philosophy-as-applied-moral-philosophy paradigm are insensitive.⁵ For example, Colin Farrelly’s (2007a) claim to be pursuing a nonideal theory approach that shares Dunn’s conception of practical political philosophy is better characterised as being what Frances Kamm (1996) calls ‘low theory’, or applied

⁵ The realist challenge is addressed by Sangiovanni (2008a).

ethics. In contrast to the high theory of utilitarianism, Kantianism or contractarianism, low theory develops theory in response to particular moral problems.

More is at stake in this debate than simply a relationship of theory to practice. What needs to be clarified is the relationship between two theories and two practices (Gunnell 1996): first, a philosophical theory that reflects on its own fact-sensitivity, its own relationship to practice, and second, everyday practices embedded within the causal order of the world but also informed to a greater or lesser extent by theoretical orientations, and manifesting a level of sensitivity to philosophical theory.⁶ Rousseau recognised the danger of simply moderating the former in the hope of exerting some influence over the latter. In rejecting the suggestion that the philosopher should propose some good that can be allied with the existing evil, he warns ‘in this alliance the good is spoiled, and the evil is not cured’ (Rousseau 1979 [1762]: 34).

A quantitative scale of fact-sensitivity obscures a qualitative change: an extreme nonideal theory would agree with Gray that ‘ethics cannot be other than an empirical inquiry’ (Gray 2000: 35).⁷ This is not a more realistic theory, but a realist argument against normative theory. In the next section I therefore turn to an examination of Williams’s defence of realism against moralism in political theory, and his associated criticisms of Rawls, as the basis for a more adequate view of what is at stake in the ideal theory debate. As has already been implied, this requires assessing contrasting views of the qualitative significance of facts to political theorising rather than simply employing a quantitative scale of fact-sensitivity.

1.2 Realism vs. Moralism

Williams identifies and criticises two species of political moralism. The first is an enactment model of the relation of morality to political practice, of which utilitarianism is the paradigmatic example. On this model, political theory adopts a panoptical perspective. It ‘formulates principles, concepts, ideals and values; and politics (so far as it does what the theory wants) seeks to express these in political action, through persuasion, the use of power and so forth’ (Williams 2005a: 1). Rawls has such an approach in mind when he contrasts his democratic view of political philosophy with Platonic views that separate the truth about justice from what is required to realize that truth (LHPP: 3). I will refer to this as the Platonic paradigm of the relationship between theory and practice. The second species of political moralism that Williams criticises is a

⁶ A double hermeneutic in Giddens’s (1993) terms.

⁷ For a critique of the positivistic social theory on which this depends, see Kelly (2006).

structural model where ‘theory lays down moral conditions of co-existence under power, conditions in which power can be justly exercised’. Williams associates this model with Kantian views and sees Rawls’s *TJ* as the foremost modern example. I will refer to it as the Kantian paradigm. Both the Platonic and Kantian paradigms are criticised by Williams for prioritising the moral over the political, seeing political philosophy as ‘something like applied morality’ (Williams 2005a: 2).

Compared to Dunn’s view this represents an advance. Williams moves beyond a dichotomy of utopianism and realism, contrasting normative principles of political legitimacy that hermeneutically ‘make sense’ within historically bound political practices with those that do not. However, Williams shares with Dunn the view that Rawls proposes a brand of ahistorical, moralising Kantianism that seeks a view of the world *sub specie aeternitatis*. It is worth recalling, however, that Rawls concludes *TJ* by claiming to have reinterpreted the point of view of eternity so that it is not a perspective from beyond the world, but rather ‘a certain form of thought *and feeling* that rational persons can adopt *within this world*’ (*TJ*: 587/514, emphases added). Rawls argues that Kant was wrong if he meant our social situation in the world to be irrelevant to the construction of first principles of justice (*TJ*: 257/226).⁸

We must not neglect the importance of Rawls’s view that political philosophy ‘has its own distinctive features and problems’, so that justice as fairness is a ‘political conception of justice for the special case of the basic structure of a modern democratic society’ (*JF*: 14). Ideal theory ‘defines a perfectly just basic structure’ (*PL*: 285). In his most revealing elaboration of the nature of ideal theory in *Justice as Fairness: A Restatement* (*JF*), Rawls says:

Discussion of this case [the nature and content of principles of justice for the basic structure of a well-ordered society] is referred to in justice as fairness as ideal, or strict compliance, theory. Strict compliance means that (nearly) everyone strictly complies with, and so abides by, the principles of justice. We ask in effect what a perfectly just, or nearly just, constitutional regime might be like, and whether it may come about and be made stable under the circumstances of justice . . . and so under realistic, though reasonably favourable, conditions. In this way, justice as fairness is realistically utopian: it probes the limits of the realistically practicable, that is, how far in our world (given its laws and tendencies) a democratic regime can attain complete realization of its appropriate political values – democratic perfection, if you like (*JF*: 13).⁹

⁸ For criticism of Rawls’s departure from Kant’s view of the relation between theory and practice, see Galston (1982); Grcic (1983); Nasr (1991).

⁹ I defend an interpretation of the fundamental continuity of Rawls’s project in Chapter 6, but such continuity is already suggested by this passage from *JF*, which is based on lectures given by Rawls in the 1980s.

A ‘perfectly just’ society is, then, a well-ordered society under the circumstances of justice that is stable for the right reasons (Freeman 2007a: 460). Ideal theory is, moreover, part of a realistically utopian conception of political philosophy that follows Rousseau in ‘taking men as they are and laws as they might be’. I will argue that Rawls develops a Rousseauian paradigm of the relationship between theory and practice. First, though, I briefly elaborate on the assumptions of the Platonic, Kantian and Rousseauian paradigms.

1.3 Realistic Utopianism

The Platonic, Kantian and Rousseauian paradigms of the relationship between philosophical theory and social practices are represented in Figure 2. Cohen is situated as an advocate and Dunn as a critic of the Platonic paradigm; Ronald Dworkin and Farrelly are situated within a Kantian paradigm; and Rawls is presented as proposing an alternative Rousseauian paradigm. These paradigms will be discussed in turn according to their conceptions of (1) reason,¹⁰ (2) the relationship between normative principles and facts, and (3) their self-understanding of how normative principles could be realized in, or are related to, practice.

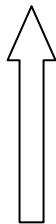
¹⁰ The following conceptions of reason within the Platonic and Kantian paradigms parallel the taxonomy of O’Neill (1996: 50):

End-oriented reasoning	Act-oriented reasoning
‘Platonist’	‘Kantian’ or Critical
<i>Instrumental</i>	<i>Particularist</i>

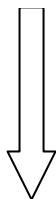
Platonic paradigm

Utopian desirability

G.A. Cohen



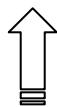
Application theory



Kantian paradigm

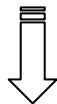
Ideal feasibility

Dworkin



Low theory

Farrelly



Real feasibility

Williams

Rousseauian paradigm

Rawls: realistically utopian realizability

Conception of justice for basic structure
(Political constructivism)



Public reason
(Reconciliation to our social world)

Rousseauian paradigm

Habermas: utopian realist realizability

Law

Reconstructive



Democracy

Realist feasibility

Dunn

Figure 2. Platonic, Kantian and Rousseauian Paradigms of the Relationship between Theory and Practice

1.3.1 The Platonic Paradigm

The Platonic paradigm is characterised by: (1) A model of ideals of intrinsic value apprehended by theoretical reason; (2) A positivistic separation of fact and value, involving an external relation between fact-independent philosophical ideals and the factual concerns of psychology and sociology; (3) A concern for the rational moral demands justice makes upon individuals in terms of personal behaviour, in addition to the role of justice in specifying institutional rules.

Cohen's view is in the tradition of political philosophy defended by Pettit (1991) that subscribes to a division of labour between the desirable ('What should we value?') and the feasible ('What can we get?'). This division is endorsed by utilitarianism, both the hedonistic utilitarianism of Henry Sidgwick or Peter Singer and the so-called ideal utilitarianism of G.E. Moore (see also Goodin 1995).¹¹ In relation to Dunn's view, Cohen's utopian concern for the pure form of justice appears not so much as at the other end of a spectrum but as a mirror image. Cohen agrees with Plato that there is such a thing as justice as such that 'transcends the facts of the world', that it is the 'self-same thing across, and independently of, history', and that 'we have to retreat to (what we consider to be) justice in its purity to figure out how to institute as much justice as possible inside the cave' (RJE: 291). Dunn may reject what he regards as the utopian liberalism of Rawls in favour of political realism, but he recognises that a Platonic search for stable and timeless foundations is not unmotivated; for him, like Cohen, it is between realism and Platonic utopianism that one must choose (Dunn 1985: 162).

Given that for Cohen political philosophy is a 'branch of philosophy, whose output is consequential for practice but not limited in significance to its consequences for practice', concerned with 'what we should think, even when what we should think makes no practical difference' (RJE: 268), one might think that Cohen's conception of justice is simply an empty ideal. However, just as Hegel remarked that Plato's *Republic* 'is essentially the embodiment of nothing other than the nature of Greek ethics' (Hegel 1991 [1820]: 20), so Cohen's defence of utopian justice can be read as an embodiment of postsocialist Marxian ethics. Cohen argues that although socialist values have 'lost their mooring' in capitalist structure (SFE: 6), this does not affect their validity.

To the extent that Marxism is still alive . . . it presents itself as a set of values and a set of designs for realizing those values. It is therefore, now, far less different than it could once advertise itself to be from the utopian socialism with which it so proudly contrasted itself (IYE: 103).

¹¹ Kukathas and Pettit (1990) interpret Rawls in this framework, but Pettit (1991) situates Rawls, and Habermas, in an intermediate category of the 'eligible'.

Cohen's project involves 'articulating and defending an ideal which moves people to action to the extent that they believe in it, an ideal which needs to be defended against capitalist ideology and capitalist ideas' (Cohen 1996a).

Fundamental socialist values which point to a form of society a hundred miles from the horizon of present possibility are needed to defend every half-mile of territory gained [from the libertarian right] and to mount an attempt to regain each bit that has been lost (Cohen 1997: 31; cf. Williams 1997).

Utopias cannot exist independently of one fundamental feasibility constraint: it must be possible to have faith in them. As Habermas puts it, for Plato '[n]othing is more practical than theory itself': knowledge promises salvation (TaJ: 278). Cohen's project is divided between a theoretical philosophical investigation into the nature of justice and a practical political defence of socialism. Within the former, his work on justice involves both a positive 'luck egalitarian' argument (Cohen 1989, 1993), and a negative argument against Rawls's fact-sensitive view of justice (IYE, RJE). At the level of philosophy, Cohen sees nothing distinctive about socialism that is not captured by taking to their logical conclusions the egalitarian commitments he sees as shared by Rawls and Dworkin, but at the level of political practice, socialism demands the realization of a principle of community (WNS). Here Cohen allies himself with the Christian view that equality requires 'not mere politics, but a moral revolution, a revolution in the human soul' (IYE: 2).

Given the utopian nature of Cohen's intrinsically desirable ideal of justice, nonideal theory in the Platonic paradigm – call it 'application theory' – is required to take a pre-given ideal and consider how individuals ought to act under nonideal circumstances of inevitable failures of compliance. The work of Liam Murphy (2003) and Adam Swift (2008, 2003a) is representative of approaches of this kind (see also Stemplowska 2008). In considering such application theory, Cohen describes Joseph Carens's (1981) utopian politico-economic model of market socialism as 'the Platonic ideal of market socialism . . . a star for socialist economists to steer by as they navigate the worldly problems of socialist design' (SFE: 264). However, at the level of realist feasibility we are also faced with how to instrumentally realize as much justice as possible under existing circumstances. For Cohen, the obstacles to this are not limitations of human nature or the operation of institutions. It is rather a tractable problem for economists of designing the machinery that would make socialism run

(WNS). In modern economic parlance, realizing justice in Plato's cave is a matter of determining the limits of the feasible set.¹²

1.3.2 The Kantian Paradigm

The Kantian paradigm is characterised by: (1) A model of principles of justice as products, or constructions, of practical reason; (2) A 'two-worlds' view consisting in an external relation between principles of right justified in accordance with facts about our nature as rational moral beings and facts about human motivations and desires that are part of the natural world; (3) A concern with applying principles of justice to the structure of law, as opposed to principles of virtue that apply to individuals within this structure of law.

Kantian positions are committed to a constraint of 'ideal feasibility' where principles for an ideal world should be consistent with the laws of the natural world and facts about rational human nature. Thomas Nagel accepts that ought implies can. 'What is right must be possible, even if our understanding of what is possible can be partly transformed by arguments about what is right' (Nagel 1991: 26). Political theory will be utopian in a pejorative sense if it 'describes a form of collective life that humans, or most humans, could not lead and could not come to be able to lead through any feasible process of social and mental development' (Nagel 1991: 6).

Taxonomies of philosophical approaches are necessarily contentious. In characterising a range of views as advocating a Kantian ideal feasibility approach, I do not mean to suggest that they endorse Kant's own substantive view, much less the specifics of Kant's methodology. Rather, there is a family resemblance between a variety of theorists who have sought to associate themselves with Kantian principles, to a greater or lesser extent, and who are commonly seen as representing a dominant Kantian strain in contemporary political philosophy. That Rawls is commonly seen as the leading representative of this approach, and is so seen by Williams, is justification enough for employing this label as a prelude to challenging prevailing assumptions about Rawls's position. Moreover, Habermas's charge that Kant subordinates democratic politics to moral principles will be important to understanding Habermas's own view in Chapter 9.

In the course of this thesis I will seek to differentiate Rawls's position from that of a number of Kantian theorists whose positions are commonly associated with it. Such theorists differ, however, in whether they defend Kantian principles on the basis of a

¹² See Swift (1999, 2003b); Goodin (2001); Cowen (2007); Brennan (2007).

Kantian methodology. Kantian constructivists such as Onora O'Neill do claim to employ a methodology that is authentically Kantian in spirit, even while rejecting Kant's metaphysics. On the other hand, theorists like Nagel, Thomas Pogge and Brian Barry ultimately offer intuitionistic justifications.¹³ Finally, Dworkin (1996) and Charles Larmore defend what might be called a 'contextual Platonism about reasons'. Larmore (2008), while upholding a Dworkinian, and by extension 'Kantian', principle of equal concern and respect, explicitly rejects the Kantian constructivist idea of self-legislation.

In the present context of my concern with ideal theory I will focus on the positions of Dworkin and Pogge. Dworkin's (2001) theory of equality of resources employs the kind of desert island thought experiment that has been taken to be definitive of ideal theory (Robeyns 2008: 344-5), while Pogge argues Rawls's theory can be used to design a blueprint of ideal institutions (Pogge 1989: 12). However, in contrast to Kant (1991 [1793]) – for whom faith for practical purposes in God and the afterlife ensures that duty acquires an object in the shape of an ideal of pure reason, and practical judgement provides the link and transition between general rules and their application – Dworkin and Pogge provide substantive ideals and integrate a consequentialist element to secure their application. Dworkin's 'ideal ideal world' of fantasy forms the basis for a theory of egalitarian improvement in the 'ideal real world', which in turn can be used to judge performance in the 'real real world'. Likewise Pogge, on the basis of a broadly consequentialist interpretation of Rawls, can argue that our responsibility to assess alternative feasible institutional schemes is in no way dependent on whether an ideal scheme is realistically attainable.

Opposing such ideal feasibility views are 'real feasibility' views that dispute whether ideal principles can gain any traction in the 'real world'. These views are united by an opposition to Kantianism and an emphasis on the priority of democracy and politics to morality. Williams offers a clear example, along with Michael Walzer (1981, 1984), Richard Rorty (1991a) and Jeremy Waldron (1999).¹⁴ The *sui generis* law theory approach advocated by Farrelly, focused on reconciling justice and virtue in the context of particular problems, can be located between these ideal feasibility and real feasibility views.¹⁵

¹³ Although Barry (1989: 271-82) refers to his position as constructivist, he sees constructivism as a heuristic for focussing intuitions. On Nagel's intuitionism, see Korsgaard (1996a: 41 n.66; 1996b). See Pogge (1981) for an appeal to intuitions and Pogge (2002) for a focus on law.

¹⁴ See also Barber (1989); Honig (1993); Wolin (1996); Sangiovanni (2008b); Galston (2009).

¹⁵ See also Sher (1997); Sreenivasan (2007).

1.3.3 The Rousseauian Paradigm

Rawls's elaboration of a Rousseauian paradigm is characterised by: (1) A model of political principles of justice as constructed out of political practical reason, practical reason that draws on ideas implicit within the domain of the political; (2) A view of political principles of justice as independent of natural facts and specific social facts but as maintaining an internal relation to general facts about the social world – a well-ordered society achieving congruence between citizens' sense of justice and conceptions of the good, uniting what right requires with what interest prescribes; (3) A concern with how a conception of justice specifies a relationship between persons within the basic structure of society, establishing a conception of public reason through which citizens exercise coercive power over one another as a collective body.

Rawls's realistically utopian approach to political philosophy draws on Rousseau's maxim of 'taking men as they are and laws as they might be'. Rawls understands taking persons' moral and psychological natures as they are to involve understanding human potentiality, or how persons' natures might be within a framework of political and social institutions organized as they should, or ought, to be (LP: 7). '[T]he limits of the possible are not given by the actual, for we can to a greater or lesser extent change political and social institutions and much else' (LP: 12; JF: 5). A constructivist procedure of probing the limits of practicable political possibility might seem to be in tension with another role Rawls allots political philosophy, that of reconciling us to our social world. But Rawls's concern with reconciliation is a concern with how the potential for political autonomy made available by our social world and identified by a constructivist procedure can be actualized in practice (Neuhouser 2000). In reconciliation by public reason we become reconciled to our social world. Constructivism seeks to elaborate a conception of social unity that could arise consistent with the historical constraints, the laws and tendencies, of our social world – or what elsewhere Rawls calls the 'circumstances of the social world' (PL: xxi) – since 'until we bring ourselves to conceive how this could happen, it can't happen' (CP: 395). Understanding the potential implicit in a public political culture for realizing an ideal is not independent of actualizing that potential. Political philosophy is realistically utopian when it 'extends what are ordinarily thought to be the limits of practicable political possibility and, *in so doing*, reconciles us to our political and social condition' (LP: 11, emphasis added).

This enables us to understand how Rawls's ideal of a well-ordered society is an ideal of democratic perfection. Ideal theory shares with Rousseau and Kant the idea of unanimity among rational persons. But in understanding the principles that are chosen in

the original position as applying first to social institutions, justice as fairness accepts ‘the social nature of the virtue of justice, its intimate connection with social practices as often noted by idealists’ (TJ: 110/95). An ideal ‘is the conception of an individual thing (a particular), determinable and even determined by an idea of reason’ (LHMP: 212). This is an employment of practical reason, which produces, or perfects, objects according to a conception of those objects, as opposed to theoretical reason which is concerned with the knowledge of given objects (PL: 93). What the circumstances of a democratic society demand on Rawls’s view is that an ideal of a perfectly just constitutional democratic regime not be deduced from apriori or perfectionist principles. Philosophy for Rawls is not prior to democracy, but rather shares a common source in shared human reason. Philosophical reflection begins within a tradition of thought, and for us within a tradition of democratic constitutional thought (LHPP: 2; see also Kaufman 2006). In ideal theory, a conception of justice is constructed by a procedure that incorporates fundamental ideas implicit in the public political culture of a democratic society. It excludes no general facts about the circumstances of justice, circumstances like the fact of reasonable pluralism that reflect the ‘historical conditions under which modern democratic societies exist’ (JF: 84). Ideal theory is theory for a society where the social order is no longer regarded as a fixed natural order (TJ: 548/480; PL: 15), and where consequently citizens possess a sense of justice, being motivated by the conception-dependent desire to act from principles that express a political ideal of citizenship (PL: 82-5). Political philosophy enters into practice through its role in educating citizens to ideal-conceptions as part of the background culture of civil society. The role that political philosophy plays is ‘not so much in day-to-day politics as in educating citizens to certain ideal conceptions of person and political society before they come to politics’ (LHPP: 7).

In contrast to the Kantian paradigm, Rawls’s Rousseauian paradigm has no place for a distinction between two worlds, the ideal categorically distinct from the real. There is therefore no question of seeking to bridge, in quasi-Platonic or quasi-Kantian fashion, a gap between an ideal perfect world and our nonideal imperfect world (cf. Phillips 1985). Like Walzer, Rawls seeks to interpret our shared understandings in a manner that democratic citizens can endorse, not to construct a Dworkinian ‘ideal ideal world’, which is then brought back down to earth (PL: 44-45). ‘Political philosophy does not . . . withdraw from society and the world. Nor does it claim to discover what is true by its own distinctive methods of reason’ (PL: 45). Hence Rawls’s description of his approach as ideal-based in contrast to Dworkin’s argument that it is rights-based (CP: 400-1 n.19). Dworkin rejects Rawls’s singular focus on the basic structure, pursuing a more ‘retail level’ approach designed to be of help to legislators and judges (Dworkin 2002). Rawls,

by contrast, sees justice as fairness as ‘addressed not so much to constitutional jurists as to citizens in a constitutional regime’ who exercise coercive power over one another (PL: 369; cf. Waldron 1999: 22). This is not the Kantian coercive power of positive law, but the Rousseauian coercive power of the public reason of citizens. In a democratic society ‘public reason is the reason of equal citizens who, as a collective body, exercise final political and coercive power over one another in enacting laws [through voting (PL: 217)] and in amending their constitution’ (PL: 214, see also 136, 139, 216, 390). Again in a Rousseauian sense, students of philosophy formulate ideas of justice not from a position of privileged philosophical insight but as citizens among other citizens (PL: 427). The ideal conception of citizenship for a constitutional democratic regime given by public reason – presenting how things might be taking people as a just and well-ordered society would encourage them to be (PL: 213) – is realized when in following the duty of civility persons act from the idea of public reason.

Rawls accepts neither a dichotomy between utopianism and realism nor one between the ideal world and the real. Facts such as the fact of reasonable pluralism affect the realizability of justice as fairness, but in thereby viewing facts in relation to human agency, Rawls seeks to identify facts that are constitutive of the ideal of justice rather than simply an obstacle to its feasibility (Bohman 2005). The justificatory relationship of reflective equilibrium between constructivism and reconciliation is circular – drawing on ideas that reflect the circumstances of our social world in order to reform the basic institutions that form that social world (Michelman 1995) – but this is a virtuous spiral of constructivist and historical social learning (Müller 2006), not a definitional vicious circle divorced from social reality.

1.4 The Pursuit of the Ideal within our Social World

Rawls’s idea of a social world is a key idea for understanding how, far from beginning *ab initio*, ‘Rawls always begins *in mediis rebus*’ (Dreben 2003: 322). Rawls takes our social world to be formed by background or basic institutions through the operation of which we become free and equal citizens. This is the background against which philosophical reflection takes place and towards the reconstruction of which it is directed. The basic structure of our social world provides the focal point of Rawls’s approach in relation to which it may be situated within social space and against historical time. Principles of justice for the basic structure form the social space within which conceptions of the good are pursued. Rawls agrees with Berlin that not all values can fit within one social world: there is no social world without loss (PL: 197-98). But

this idea of social space is not a given empirical fact; it can only be identified from the standpoint of a normative conception of a well-ordered society. The pursuit of the ideal is not, *pace* Berlin, the pursuit of a Platonic utopia: accepting the circumstances of our social world, in particular the fact of reasonable pluralism, ideal theory is an ideal theory of conflict not a theory of nonconflict (Dreben 2003: 323). An ideal form for the basic structure is a framework that provides a principled basis for prioritising among competing values guided by the priority of the right. We ‘strive for the best we can attain within the scope the [social] world allows’ (PL: 88).

Justice as fairness starts from domestic justice (principles applying to the basic structure of a closed society), works outwards to global justice (constructing principles of a Law of Peoples for a Society of Peoples), and inwards to local justice (principles applying directly to institutions and associations within the basic structure) (PL: 259-62; JF: 11). Although as a privileged locus for our considered judgments the sphere of domestic justice constitutes our starting point, ultimately justification is a matter of reflective equilibrium between all these levels. An account of what is required by global justice may require us to revise our account of domestic justice, while principles for the basic structure constrain, but do not determine, principles of local justice, in Jon Elster’s (1992) sense, and are characterised sufficiently loosely to allow adjustment to future conditions and changing social circumstances (JF: 11-12). Local justice focuses on the mechanisms through which principles of justice are realized in particular institutional situations.

Second, we can locate Rawls’s approach within a temporal perspective. While the politician looks to the next election and the statesman to the next generation, philosophy looks to ‘the indefinite future’. The student of philosophy looks to the ‘standing conditions of human life’, the ‘permanent conditions and the real interests of a just and good democratic society’ (CP: 447, 492; LP: 97). Therefore, in contrast to a Kantian two-worlds view, Rawls’s view involves nested spheres, the standpoints within them recalling a Russian doll and associated with different temporal perspectives.

So far I have argued that the basic structure of our social world forms the focal point of Rawls’s conception of justice. Questions of local justice pursued by the social scientist concern how institutions and associations operate within the basic structure, while statesmen and politicians work within the framework of a conception of justice for the basic structure. This provides a framework within which we may understand the relationship between ideal theory and nonideal theory in Rawls’s realistically utopian political philosophy.¹⁶ As is particularly evident in *The Law of Peoples* (LP), ideal

¹⁶ My understanding of Rawls’s ideal theory follows Schwarzenbach (1990).

theory consists of two parts: first, the construction of a conception of justice for a well-ordered society (political constructivism), and second the consideration of whether this could come about and be made stable under the circumstances of our domestic and international social worlds (reconciliation to our social world).¹⁷ While initially simply posited, strict compliance relates to Rawls's conception of stability; it is thus a normative requirement of a fully justified conception of justice, whose possibility requires defence, not an unvindicated theoretical simplification. The problems addressed in nonideal theory concern partial compliance and unfavourable conditions. The examples of noncompliance that Rawls discusses in most detail, civil disobedience and conscientious refusal, must be carefully distinguished from acts of simple law-breaking. Rawls assumes a state of near justice in which civil disobedience is tolerable, because potentially justified, noncompliance. Noncompliance may succeed in showing that laws do not comply with principles of justice. Indeed, 'A general disposition to engage in justified civil disobedience introduces stability into a well-ordered society, or one that is nearly just' (TJ: 383/336).

Rawls's principles define a 'perfectly just society, given favorable conditions', conditions, that is, that allow for the priority of liberty (TJ: 351/309). This should not be taken to mean that they are principles for an ideal world that is not our own. Ideal theory is both historically and contextually indexed: lesser political liberty may have been justified in today's liberal societies under some historical situations, and it may be necessary to forgo some political liberties in order for existing burdened societies to develop into ones in which all the basic liberties can be fully enjoyed (TJ: 247/217). However, Rawls argues that in the United States, and presumably in other liberal societies, reasonably favourable conditions for the priority of liberty *do* obtain, even if the political will to achieve a well-ordered society may not (PL: 297).¹⁸ Crucial to this is the fact that Rawls takes it that citizens do not view the social order as a fixed natural order. *Historical limitations* that in other times meant, or in other places mean, that favourable conditions did not or do not obtain must be clearly distinguished from *natural limitations*, necessitating, for example, the regulation of liberty for public order, that obtain even in the most favourable historical conditions (TJ: 245/215). While the former concerns *unequal* liberty, the latter concerns *less extensive* liberty.

¹⁷ For the contrary view that ideal theory applies only to Rawls's principles of justice, see Phillips (1985: 568 n.4).

¹⁸ I leave aside the complications introduced by Rawls's distinction in the original edition of *TJ* between a general conception of justice and a special conception suited to favourable circumstances conducive to the priority of liberty as the general conception is largely omitted from the revised edition.

The structure of *LP* helps to clarify the roles of ideal and nonideal theory in Rawls's other major works, but *LP* also helps to clarify upon whom the responsibilities to respond to the problems of nonideal theory fall. The student of philosophy looks to the real interests of a well-ordered society and, on this basis, first sets out a conception of justice in ideal theory, and second says enough to show how the principles of ideal theory provide guidelines that can and should inform how we approach the problems of nonideal theory. However, it is the task of the statesman, not the student of philosophy, to 'discern these conditions and interests in practice' and to convince public opinion of the importance of the duty of following the doctrine of just war and of the duty of assistance to burdened societies (*LP*: 97). In like manner, civilly disobedient citizens seek to discern these interests in practice by appealing to their fellow citizens' sense of justice. When it comes to the theory of punishment, on the one hand, given the normal conditions of human life, '*ideal theory* requires an account of penal sanctions as a stabilizing device' (TJ: 241/212, emphasis added).¹⁹ Strict compliance theory is not for a society of moral saints. On the other hand, ideal theory seeks only to provide guidelines for how the partial compliance theory problem of punishment is to be resolved in practice. It is limited to consideration of principles that justify sanctions according to the principle of liberty and its priority. Finally, with respect to restricting liberty when faced, for example, with rival religious sects, all arguments must similarly proceed from the ideal theory assumption of the priority of liberty: to this degree the priority of liberty carries over to partial compliance theory (TJ: 242/213).

While ideal theory sets out guidelines for addressing the problems of nonideal theory, addressing these problems ultimately involves political practitioners making *judgements in practice* regarding how particular problems are to be decided in accordance with the framework provided by ideal principles. The rules with which we are concerned in ideal theory are constitutive rules that 'construct actions-as-realizations-of-ideals' rather than instrumental rules that identify 'preexisting actions that serve as a means of promoting goals'. On the Platonic paradigm

the actions governed by moral rules have intrinsic value, but they are constituted independently of those rules. The function of the rules is to detect the actions that bear intrinsic value, and the point of acting on those rules is, in some (usually unspecified) sense, to 'conform' to that value (Schapiro 2006: 57).

¹⁹ Cf. Mandle's (1997) interpretation of the ideal/nonideal theory distinction as a distinction between distributive and retributive justice.

By contrast, on the Rousseauian paradigm, ‘moral rules constitute the actions they govern, so as to make them fit to fulfil a function in which their intrinsic value consists’ (Schapiro 2006: 57).

It may be, as Rawls says, that nonideal theory addresses the pressing and urgent problems that we are faced with in everyday life. But on Rawls’s conception of political philosophy that is where they ought to be tackled: in everyday life (or at least in disciplines such as social science that are concerned with the immediate problems of everyday life). The competence of philosophy to address these problems is limited (Freeman 2007a: 232). Principles of justice in ideal theory specify a structural ideal to guide political action (PL: 284-5). This is a ‘social ideal by reference to which political judgment among second-best, and the long-range direction of reform, may be guided’, guided by theory, but ultimately reliant on good sense and plain hunch (CP: 150; TJ: 278/246).

Rawls reverses the relationship between theory and intuitions present in utilitarianism. Whereas utilitarians often begin from intuitions about the good and then proceed to consider how the good is to be realized and trade-offs made, Rawls provides a framework of practical principles that can guide judgements in practice. Our natural duty to remove injustices depends upon intuitive judgements regarding departures from an ideal of justice, but this is guided by the lexical order of the principles of justice. ‘Thus while the principles of justice belong to the theory of an ideal state of affairs, they are generally relevant’ (TJ: 246/216). Rawls argues, against Sidgwick, that we should pursue a theory of ideal social relations rather than beginning from existing circumstances and the here and now, and he furthermore denies the conclusion that Sidgwick draws that this makes political philosophy an inquiry into the ideal society that bears no relation to the practical problems of actual life (Sidgwick 1981 [1907]: 15-22). In the process, the ambitions of political philosophy are both reduced and extended. They are reduced inasmuch as, to be consistent with the priority of liberty in a democratic society, the focus of political philosophy is restricted to public principles for an ideal form of the basic structure and does not include direct policy prescriptions. However, they are extended to the extent that ‘[t]he primacy of justice is inseparable from the empowerment of citizens’ (Audard 2007: 72). Rawls’s realistic utopianism is consistent and attractive. It argues for an important but abstract role for political philosophy, indeed a role that is important because abstract (J. Cohen 2004), while evincing sensitivity to the realist criticisms of Dunn, Williams and Berlin (see Hanley 2004) that political philosophy cannot, and should not aspire to, usurp political judgement.

It is, though, a long-running theme in criticisms of Rawls that this is insufficient given the nonideal world in which we live (Feinberg 1973; RJE: 4-5). From the perspective of history, social science and utilitarian moral philosophy, our world is indeed imperfect: actions have unforeseen consequences, causal connections are complex and people fail to comply with principles of justice. But in view of the currency that the idea that we live in a nonideal world has obtained in Rawlsian political philosophy, it is worth considering what this might mean in Rawls's philosophical terms.²⁰ The closest Rawls comes to describing nonideal theory as concerned with a nonideal world is when he describes its concerns as 'questions arising from the highly nonideal conditions of our world with its great injustices and widespread social evils' (LP: 89). Notice, however, that it is *conditions* that are nonideal and that evils are *social*. The nonideal conditions of the world are only identified as such relative to a realistically utopian idea of a well-ordered society that provides grounds for reasonable faith that these social conditions can be overcome.

At its most banal, to say that we live in a nonideal world might be to point out, as Farrelly does, the many practical problems involved in judging how a philosophical ideal applies in practice. I have argued, however, that this overlooks how for Rawls the problems of nonideal theory are only identified with reference to an ideal and therefore mistakes the level at which Rawls argues political philosophy should operate. At the other extreme, to say that our world is nonideal might be to assume a standpoint similar to Leibniz's ethics of creation, but taking the opposite view to his that God creates the best, the most perfect, of all possible worlds (TJ: 159/137; LHMP: 109). However, Rawls is concerned to deny the existence of the perfectionist standpoint that would be required to make such a judgment of imperfection. Alternatively, in a weaker form, and in a secular analogue of original sin, one might think that human nature under capitalism is so corrupted, and fear and greed so prevalent, that the achievement of a just society is beyond us. Our world would be nonideal in the sense that there is no foreseeable prospect of escape from the nonideal circumstances we face. Cohen may sympathise with this view, but it runs contrary to Rawls's belief that, as Rousseau put it, 'man is naturally good, and that it is solely by [our] institutions that men become wicked' (Rousseau, quoted in J. Cohen 2002).²¹

²⁰ A contrast between ideal theory and the nonideal world is made by Feinberg (1973); Pogge (1989: 136); Gray (1991: 191); Brennan and Pettit (2005); Evans (2005); Screenivasan (2007); Wenar (2008).

²¹ The 'center and pivot' of Rousseau's thought is the 'decisiveness with which he rejected every thought of the *original sin* of man' (Cassirer 1954: 73, 74). This sentiment, and the way in which it opposes an Augustinian view of original sin, is discussed by Rawls in LHPP: 204-8. See also J. Cohen (1997a); Freeman (2007a: 19). Cf. Cohen: 'Is it an axiom that human beings are *capable* of justice? Is "original sin" a contradiction in terms?' (RJE: 330). For Cohen's reasons for

Rawls is ultimately concerned with the idea of a well-ordered society because of a concern with whether human life is redeemable.²² His conception of justice is not intended to provide ideal policy prescriptions, but rather to guide our actions by shaping the attitudes we bring to politics. Given reasonably favourable conditions, conditions which Rawls believes do obtain, justice as fairness ‘understands itself as the defense of the possibility of a just constitutional democratic regime’ (PL: 101). For Rawls, if human nature is such that a well-ordered domestic and international society is not realizable, because human beings are largely amoral, ‘one might ask, with Kant, whether it is worthwhile for human beings to live on the earth’ (LP: 128; see also PL: lxii).²³

1.5 From Realistic Utopianism to Utopian Realism

Samuel Freeman has highlighted how the idea of a perfectly just well-ordered society is the central theme of Rawls’s work, and how Rawls’s concern with whether this idea is realistic is influenced by Rawls’s attachment to an ideal of a democratic society that started with Rousseau (Freeman 2007a: 461). But Freeman observes that by the end of his career Rawls had come to see general agreement on justice as fairness as not realistically possible. Rawls says, ‘while realization is, of course, not unimportant, I believe that the very possibility of such a social order can itself reconcile us to the social world’. Here Rawls risks succumbing to an impotent idealism; but as he goes on to say, the justification for this view rests on the fact that an ideal of a well-ordered society is ‘not a mere logical possibility, but one that connects with the deep tendencies and inclinations of the social world’ (LP: 128).

Two avenues are open to post-Rawlsian political philosophy. First, we might maintain that Rawls articulated, or aspired to articulate, a true and ‘timeless snapshot of an ideal world’ (Jensen 2009: 184). One might then, with Farrelly, follow a meliorist low theory approach which accepts that this ideal is beyond the present horizon of possibility, but, in focusing on practical problems, aspires to ‘help revive the transformative potential of liberal democracy’, which ‘has been stifled by the principled paradigm of ideal theory’ (Farrelly 2007a: 232). Or, also following this first view, one might follow Stears’s insightful observation that the reason Rawls believes that radical political action is not required in current circumstances is not because of theoretical

abandoning the view that human nature is plastic, and consequent acceptance of a sociological view of ineradicable human selfishness analogous to original sin, see IYE: 118-20.

²² Neiman (2002: 310-14); Müller (2006: 336); Pogge (2007: 26); Freeman (2007b: 325); Weithman (2009).

²³ Cassirer (1954: 70) calls this a ‘genuinely Rousseauist thought and sentiment’.

abstraction from the real world, but because Rawls, and liberals generally, ‘believe that their goals are imminently achievable, even *immanent* in the system as it currently exists’ (Stears in King 2008: 238).²⁴ One might then condemn Rawls’s idea that philosophy forms part of the background culture of civil society as ‘like saying its goal is hanging the wallpaper in Plato’s cave’ (Smith 2007). Explicitly rejecting Rawls’s democratic view of political philosophy in favour of a radical application theory, one would have to counter Rawls’s view that the alternative to a democratic view is a Platonic view that separates the question of the truth about justice from what is required to realize that truth within the world. However, in misconceiving of the value of philosophy – the latter approach in a way that is more self-delusional than dangerous – it is likely that only disillusionment will follow. As an alternative, one might take seriously Rawls’s thought, following R.G. Collingwood, that the history of political philosophy is the ‘history of a problem more or less constantly changing, whose solution was changing with it’ (LHPP: xiii, 103). We have to concede that as established beliefs and social theory change, ‘it is possible that the principles of justice which it seems rational to acknowledge may likewise change’ (TJ: 548/480).

We should seek a post-Rawlsian approach that neither deflates the ambitions of political philosophy in advocating a low theory project of piecemeal policy advice, nor runs the risk of deluded self-importance inherent in a radical application theory. Williams recognises that the circularity present in a holistic constructivist model of justification such as that employed by Rawls is no objection to it as a normative theory, so long as the set of concepts involved has a grounding in reality and is not forced upon a recalcitrant world (Williams 2005b: 21, 23). Where Williams’s criticisms of Rawls do ultimately hit home is in the observation that ‘no political theory, liberal or other, can determine by itself its own application’. Any utopian political theory, even a constructivist one

will seem to make sense, and will to some degree reorganize political thought and action, only by virtue of the historical situation in which it is presented, and its relation to that historical situation cannot fully be theorized or captured in reflection (Williams 2005b: 28, 25).

The limitation of Rawls’s approach is that it can only understand its relation to history from within its own framework. It can reflect on the changing way the tradition of political philosophy should respond to political problems, but not on the place and role of philosophy within the social world. It is therefore insensitive to the fact that

²⁴ Stears here draws upon Waldron (1993). See also Stears (2007).

philosophy can no longer rely on the background of a shared social world and associated traditions of thought. Williams remarks that analytic philosophy has been notably ill-equipped for ‘reflexively raising questions of its own relation to social reality’ (Williams 2006: 159). I have sought to partially defend Rawls on this front, but it is at this point that the limitations of Rawls’s approach are reached.

Williams endorses Habermas’s aspiration to situate philosophy between facts and norms, and I will argue that pursuing Habermas’s approach is true to what the spirit of Rawls’s approach demands under changed social conditions. There is a tension between the form of Rawls’s approach – its understanding of the realistically utopian role of political philosophy in which it probes the limits of practicable political possibility – and its content, which takes the form of principles of justice that are the most important part of an ideal of a realistic utopia. In Habermas’s terms, it is a tension between the procedural and substantive aspects of Rawls’s approach. However, Williams concedes more of the utopian dimension of political philosophy than does Habermas. Williams does not attend to the importance of Habermas’s paradigm shift from the philosophy of the subject to an intersubjective philosophy of language. The postmetaphysical proceduralist concept of rationality that results ‘cannot sustain utopian projects for concrete forms of life as a whole’ (RR: 87). While philosophy becomes more modest in this sense, this modesty is linked with the ambition of preserving a view of the utopian dimension of practical reason. Philosophy as reconstructive critical social theory seeks to reconstruct the tension between facticity and validity present in moral discourse and the law. In this sense, if Rawls is a realistic utopian, Habermas is a utopian realist.²⁵

Habermas’s elaboration of a Rousseauian paradigm provides a final paradigm of the relationship between theory and practice (Table 1). His utopian realist position can be seen as involving: (1) A reconstructive theory of communicative reason that identifies the ideal presuppositions that are and must be appealed to in discourse; (2) A reconstructive theory of the tension between facticity and validity, identifying how action orientated in accordance with context-transcending ideals is stabilised within social order; (3) A conception of transcendence from within such that principles of the right are appealed to in order to compensate for the breakdown of shared conceptions of the good and become the basis of social order, a conception within which law plays a crucial role in the social mediation between facts and norms.

²⁵ I borrow this suggestion from Rosen (1999). See also Giddens (1991: 154-8).

	Platonic paradigm: utopian desirability	Kantian paradigm: ideal feasibility	Rousseauian paradigm	
			Rawls: realistically utopian realizability	Habermas: utopian realist realizability
Conception of reason	Theoretical reason	Practical reason	Political practical reason	Communicative reason
Relation between normative principles and facts	External relation between ideals of justice/the good and facts	External relation between principles of right and facts of sensible world	Internal relation between political conception of justice and general facts of social world	Internal relation between validity and facticity
Route of realization	Law and individual behaviour	Law and virtue	Conception of justice for basic structure and public reason	Law and democracy

Table 1. Reason, Facts and Realization in the Platonic, Kantian and Rousseauian Paradigms

In reflecting on Rawls's legacy, David Estlund (2003a) refers to Rawls's 'audacious humility'. Guided by this idea, one should see Rawls's reluctance to offer practical policy prescriptions not as humble conservatism, but as an audacious claim to have identified the potential for the realization of a well-ordered society immanent within, and realizable through reforming, the 'groundwork of our existence' (PL: 139, quoting J.S. Mill). While I have not endorsed the substance of this view, I have argued that we have much to learn from its methodological approach. It is only by conceiving of political philosophy on the model of applied moral philosophy that one can be led to the view that the 'beauty and power of Rawls's conception are wasted insofar as they cannot be *imported* into the political arena' (Pogge 1989: 6, emphasis added). One is only following Rawls's own view when one recognises that it is a 'great mistake' to 'see a political conception of justice as something that will tell [one] what to think', to think that there is a theoretical device that will 'deliver the answers to all sorts of questions when you want them' (Rawls 1991: 45).

2. Post-Kantian Constructivism and the Problem of Self-grounding

Having outlined a methodological framework for analysing the relationship between theory and practice, in the remainder of this thesis I seek to chart a course for post-Rawlsian political philosophy by analysing Rawls's constructivist methodology in relation to its two most prominent critics, Cohen and Habermas. At an interpretative level, I will argue that Rawls's approach has greater affinities with Habermasian critical theory than it does with Cohenite analytic philosophy. At the level of normative advocacy, I defend Rawls against Cohen's criticisms, and a number of Habermas's, but conclude that Habermas ultimately both identifies fundamental weaknesses in Rawls's approach and proposes a superior alternative paradigm for political philosophy.

A comparison of Cohen, Rawls and Habermas is immediately complicated by the existence of two cross-cutting cleavages. The first cleavage sets the political and philosophical pragmatism of Rawls and Habermas against Cohen's continued allegiance to a Marxian ideal of equality, Perry Anderson locating Rawls and Habermas on the 'adjustable centre' of the political spectrum in contrast to the orthodox Marxism of the 'vanquished left' (Anderson 2005). The second cleavage sees Cohen and Habermas seemingly making common cause against Rawls, Freeman observing that both

Jürgen Habermas and G.A. Cohen object to Rawls's appeal to facts about human beings and social cooperation to justify his principles of justice. Their claim (roughly) is that fundamental moral principles are not to be grounded in the factual contingencies of the world (Freeman 2007a: 252; see also Freeman 2007b: 11, 99).²⁶

How could Habermas have such a different understanding of philosophy to Cohen, but join with Cohen in making similar criticisms of the fact-sensitivity of Rawls's position? I will argue that the Rawls-Habermas debate sowed the seeds of confusion due to the way Habermas mixed criticisms from the perspective of political theory with criticisms from the perspective of moral theory. In *Between Facts and Norms* (BFN), by contrast, Habermas seeks to mediate the abstract normativity of Rawls's philosophy of justice with the realism of the sociology of law. Far from rejecting the importance of facts to political philosophy, Habermas makes the different criticism that Rawls blurs the distinction between de facto acceptance and rational acceptability, insufficiently acknowledging the way facts inhibit the realization of principles of justice as well as

²⁶ See Rawls's acknowledgement of the shared misgivings of Cohen and Habermas with respect to the idea of an overlapping consensus (CP: 473 n.3).

serving as their basis (see Bohman 1996: 250-51 n.17). When we reach Habermas in Chapter 9, the argument of this thesis will have come full circle. Rawls seeks to overcome but does not fully escape the dualism of facts and principles that Cohen defends, but, on the basis of a paradigm shift to the philosophy of language, Habermas can overcome this methodological dualism without levelling the tension between social facts and context-transcending principles.

This chapter adds a genealogical background to the methodological framework already introduced, elaborating on what it means to describe Rawls and Habermas as developing alternative approaches within a Rousseauian paradigm, and contrasting these attempts to reconcile realism and utopianism with Cohen's progression from scientific to utopian socialism. A post-Kantian interpretation of Hegel, associated in particular with Robert Pippin and Frederick Neuhouser, provides the guiding thread of this narrative. Following Pippin, I will trace the development of a modern rational agency tradition, concerned with freedom and the deliberative capacity of practical reason through which freedom can be realized, extending through Rousseau, Kant, Fichte, Hegel, Marx, Rawls and Habermas (Pippin 2008: 21; see also Baynes 2007). Cohen, Rawls and Habermas all develop their approaches through reconstructing with the tools of analytic and post-analytic philosophy the tradition of classical German philosophy, a tradition concerned with the idea of freedom as self-legislation that begins with Rousseau. With Hegel, the problem of self-legislation takes a historical and social turn, the philosophical problem of self-grounding in modernity requiring normative foundations to be found in existing social practices. But while Rawls and Habermas engage with this problem, Cohen rejects it as a Hegelian legacy in Marxism from which the defence of a Marxian ideal of justice must liberate itself.

2.1 Rousseau and the Counterfactual Imagination

Rawls's realistic utopianism and Habermas's utopian realism may seem philosophical versions of a 'banal everyday inclination: to have one's cake and eat it' (Anderson 2005: 165). On this view, the impossible dream that haunts both *Political Liberalism (PL)* and *BFN* is of an 'unconforming conformity: the dream of extracting a radical alternative to our existing social world from the lineaments of its own description of itself' (2005: 111). Rawls is unwilling to be pinned down, 'swaying back and forth between impartial principles, tender scruple and brute fact', both erecting a utopian ideal and arguing that the realistic course of the world is 'an increasingly hopeful pilgrimage towards it' (2005: 165). Similarly, Habermas's characteristic thought pattern is 'the continual movement of

a theoretical shuttlecock, from the optative to the indicative and back again', resulting in a theory that 'answers to the responsibility neither of an accurate description of the real world, nor of critical proposals for a better one' (2005: 127-28; see also Cook (2001)).

In contrast to these interpretative approaches, Cohen's move from descriptive historical materialism to prescriptive normative philosophy implicitly rejects the 'have one's cake and eat it too' strategy of combining social theory and moral theory (Callinicos 2001). But the problem of how to have one's cake and eat it is the problem of self-grounding that faces all critical theories:

On the one hand, if the theories are too descriptive and draw their conceptual resources from the ideas and institutions which currently exist, they lose their critical force . . . On the other hand, if they articulate forms of critique that are entirely external to the systems of knowledge and practice that exist, then they threaten to become elitist and dependent on 'transcendental' support, hence lacking any basis in the experience of the subjects for whom the theory is articulated in the first place (Power 1998: 207).

The deep structure of critical theorizing consists in a counterfactual imagination that combines 'reconstructive-transcendental' and 'critical-reflective' elements. For Rawls, an Archimedean point for assessing our social world is a standpoint neither 'mired in the actual interests and desires of people in a specific setting, thereby avoiding the Scylla of infection by existing institutions', nor 'based on a priori or perfectionist premises, thus avoiding the Charybdis of unacceptable generality and abstraction' (Daniels 1996b: 152; cf. Sandel 1996: 17). With his conception of 'transcendence from within', indeed in his philosophy as a whole, Habermas too seeks to chart a course between realism and utopianism (Johnson 2005), steering

between the Scylla of a levelling, transcendence-less empiricism and the Charybdis of a high-flying idealism that glorifies transcendence . . . [T]he theory of communicative action integrates the transcendental tension between the intelligible and the world of appearances in communicative everyday praxis, yet does not thereby level it out (TT: 91).

The idea of the counterfactual imagination, or the vocabulary of the 'as-if', can be traced back to Rousseau. Rousseau's declaration that 'Man is born free; and everywhere he is in chains' represents the birth of the tradition of critical social theory (Cooke 2006: 1; Honneth 2007: 5). His *Discourse on the Origin of Inequality* offers a pessimistic description of the corrupting effect of society, but the plasticity of human nature analysed in *Emile* provides grounds for believing that these effects can be overcome through the restructuring of social institutions outlined more fully in the *Social Contract*.

With Rousseau there originates a distinction between the natural world and the social world and the consequent potential of achieving freedom in and through society.

Two issues that will feature prominently in subsequent chapters – social theodicy and the relationship between rights and popular sovereignty – have Rousseauian roots. Rousseau's solution to the problem of inflamed *amour propre* is to seek its cultivation not its suppression: the malady must be cured with its own resources. This is Rousseau's theodicy of self-love, his 'insistence on finding in evil itself a promise of redemption' (Neuhouser 2008: 266). The idea that man is born free is closely connected to the idea that man is born good,²⁷ and it is Rousseau who provides the link between the young Rawls's preoccupation with the sin of pride and the promise of redemption through community and the mature Rawls's non-theistic (but not secular) recourse to the social contract tradition. In his undergraduate thesis Rawls argued that 'any contract theory of society is false' as persons only come to exist through and in community (MSF: 126). It was only after reading Rousseau that Rawls changed his view.²⁸ A similar attempt to translate theistic conceptions of redemption and transcendence into non-theistic terms is apparent in Habermas's idea of transcendence from within. The political theory of *BFN* appeals to a communicative conception of reason that is, as the title of a subsequent book puts it, between naturalism and religion (BNR).

Locating the origins of critical theory in Rousseau suggests that critical theory has not just a narrow sense, referring to the work of the Western European Marxist tradition known as the Frankfurt School (Jay 1996), but also a broad sense, where a critical theory 'provides the descriptive and normative bases for social inquiry aimed at decreasing domination and increasing freedom', that unites the work of Rawls and Habermas (Bohman 2005). Rousseauian ideas are taken up by Kant and Hegel and in their turn Rawls and Habermas draw on Kantian and Hegelian ideas in assessing the continuing relevance of Rousseau's political ideal of a self-governing democratic community. The methodology of the counterfactual imagination is the corollary of a practical commitment to freedom as self-government. For Rousseau, the fundamental problem of politics is to find a form of association in which each individual 'while uniting himself with the others, obeys no one but himself and remains as free as before'. In losing natural freedom but gaining moral freedom, 'obedience to a law one prescribes to oneself is freedom' (Rousseau 1997 [1762]: 60, 65). This is autonomy: self-government, or, more literally, giving a law to oneself.

²⁷ 'Everything is good as it leaves the hands of the Author of things; everything degenerates in the hands of man' (Rousseau 1979 [1762]: 37).

²⁸ Rawls admits he did not 'read Rousseau with understanding' until the mid-to-late-1950s (Reidy 2009).

Within the modern polity the problem arises of the relationship between popular sovereignty and individual rights, or public and private autonomy. Habermas's idea of the co-originality of rights and popular sovereignty is arguably implicit in the work of Rousseau and Rawls (Brettschneider 2007: 3). But Habermas sees Rousseau's central innovation as a procedural understanding of legitimacy, later worked out more fully by Kant. According to Habermas, Rousseau confused the introduction of a new procedural principle of legitimization with substantive proposals for institutionalizing just rule (CES: 185-6). Procedural legitimacy is for Habermas a reflexive level of justification whose validity is independent of specific concrete institutional forms. This is certainly controversial as an interpretation of Rousseau (J. Cohen 1986a; 1997: 89 n.27; Blaug 1999: 62-8), as is Habermas's claims that 'Rousseau wanted democracy without public debate', the general will being 'more a consensus of hearts than of arguments' (STPS: 99, 98).²⁹ But in this thesis I will seek to defend Habermas's procedural elaboration of a Rousseauian paradigm against Rawls's substantive elaboration.

2.2 Kant and Self-legislation

Kant is the pivotal theorist in the freedom as autonomy tradition that begins with Rousseau. For Kant, the will is 'not merely subject to the [moral] law, but is so subject that it must be considered as also making the law for itself and precisely on this account as first of all subject to the law (of which it can regard itself as the author)' (Kant 1991 [1785]: 93). According to Rawls, 'Kant's main aim is to deepen and to justify Rousseau's idea that liberty is acting in accordance with a law that we give to ourselves' (TJ: 256/225; see also 264/233). But seeing Kant as the progenitor of an Enlightenment humanist conception of democratic autonomy is controversial and requires briefly entering a three-cornered interpretative dispute between what Karl Ameriks (2000) refers to as pure Kantians, post-Kantians and traditional Kantians. Pure Kantians offer a view of Kant purified of supposed out-of-date or inessential metaphysical elements. Influenced by Rawls, O'Neill, Nagel and Pogge can be placed in this category, whose leading exponent is perhaps Christine Korsgaard.³⁰ Post-Kantians, on the other hand, such as Habermas, Pippin and Rorty, focus on how the idea of autonomy as self-legislation was taken up by philosophers after Kant, in particular Hegel. In their joint opposition to nonsecular values and metaphysics, and consequent understanding of

²⁹ '[T]he feeling of justice stands for [Rousseau] as the true title of nobility which nature has inscribed in the hearts of men' (Cassirer 1945: 55). Compare Rawls: 'Purity of heart, if one could attain it, would be to see clearly and to act with grace and self-command from th[e] point of view' of the original position (TJ: 587/514).

³⁰ On Rawls's influence on Kant scholarship, see González (2005); O'Neill (2000: 65-81).

justification through internal procedures of reflective endorsement or reflective equilibrium, the two schools reinforce one another.

From a more traditional Kantian standpoint, Ameriks sets himself against both pure and post-Kantians, pointing out that while for Kant the moral law is not external to our essential nature, our rationality is not only *in us*; it applies for any rational being and not just human selves (see also Kain 1999, 2004). It was Kant's successors, first Reinhold and then Fichte and Hegel, who absolutized Kant's notion of autonomy in order to 'complete' Kant's Copernican revolution in philosophy. On Ameriks's opposing view, Kant's notion of autonomy is metaphysical, but moderate rather than absolute. It is metaphysical in asserting the existence of absolute freedom, but moderate in not demanding, as Kant's successors did, that this freedom be given an absolute foundation in subjectivity. As an interpretation of the historical Kant, Ameriks's traditional metaphysical view is compelling, but it is not an approach I will follow. As a practical proposition, Ameriks's call of 'back to Kant' has significant drawbacks. Ameriks risks making Kant unintelligible outside a theistic framework and unrecognisable as the first 'philosophical modernist' (Pippin 1999). The more hermeneutical appropriation of Kant I will follow involves asking, with Rawls, what Kant means *for us*.³¹ We should indeed first seek to think ourselves into a theorist's scheme of thought, their political and social world, and understand *their* questions, but we do so from our own standpoint and ultimately with an interest in how this scheme of thought informs *our* questions (LHPP: 103-104).

However, if the traditional Kantian position is unpersuasive, so too is the pure Kantian position. I focus here on Korsgaard's influential account of how normative questions arise from the reflective nature of human beings. Drawing initially upon Hume, Korsgaard (1996a) develops a view of normativity as reflexivity. Morality on this view is grounded in human nature, for there can be no standpoint beyond our own normative point of view if morality is to make claims on us. Methodologically, this means that a clear statement of the problem of reflection leads to a statement of the solution in terms of reflection. Korsgaard appeals to a notion of practical identity (a description under which an agent values themselves) to derive Kantian conclusions from Humean premises. It is our practical identity as a citizen of the Kingdom of Ends that leads to a conception of the right.

It is interesting to find Cohen criticising Korsgaard's constructivism in similar terms to Ameriks. While for Kant the moral law comes from an individual's own will, Cohen asks whether it comes *just* from the individual's own will. Can Korsgaard

³¹ Rawls's CI-procedure interprets Kant's categorical imperative 'for us' (LHMP: 210).

maintain the categorical status of morality while humanizing morality's source? Cohen argues in the negative. If I make the moral law, and this law has its basis in human nature, then I can just as easily unmake it. Cohen concludes that 'something transcending human will must figure in morality if it is to have an apodictic character . . . [I]f morality is merely human, then it is optional, as far as rationality is concerned' (Cohen 1996b: 188). Concern about whether Korsgaard can maintain the categorical status of morality is also voiced by Larmore (2008) and Habermas (IO: 32, 273 n.47).

In its focus on the individual, Korsgaard's constructivism is Fichtean in nature (Wood 1999; Pippin 2008: 88 n.47).³² Certainly Korsgaard's social conception of practical identity holds out the possibility of a rapprochement with Hegelian views (Neuhouser 2000: 328 n.32; Laden 2001), but while relinquishing the metaphysical background of Kant's view, Korsgaard does not avail herself of the resources of a Hegelian position. In responding to Williams's (1981) Humean scepticism regarding external reasons, Korsgaard makes no recourse to Rawls's reply that Kantian conception-dependent desires that blur the boundary between Humean and Kantian accounts of moral motivation, are learnt from an institutionally shaped public political culture. Given these differences between Korsgaard and Rawls, it remains to be seen whether Rawls's constructivism is equally vulnerable to criticism. While Ameriks sees Rawls as the founder of a pure Kantian school, I will argue that Rawls's position is both intended as, and best interpreted as, a *post*-Kantian view.

Although I have not sought to refute the pure Kantian position, merely to establish the parameters of the analysis to follow, I have given initial grounds for believing that pure Kantians propose an untenable middle ground; the real choice is between Ameriks's traditional Kantianism and post-Kantianism. To establish the tenability of a post-Kantian approach, however, we will need to consider whether Hegelian bootstrapping conceptions of self-legislation can consist in more than, as Ameriks puts it following Kierkegaard, 'flogging oneself to mimic real authority' (Ameriks 2000: 14).

2.3 Pragmatism and Immanent Foundationalism

Before examining post-Kantian interpretations of how Hegel takes forward the rational agency tradition, it will be instructive to consider how the development of such interpretations of Hegel, influenced by American pragmatism and particularly Wilfrid Sellars, has paralleled the increasing prominence of postfoundationalist Hegelian themes in analytic philosophy (Redding 2007). Pragmatism, I will argue, is the 'fermenting

³² For a Fichtean critique of Rawls, see Baur (2002).

agent . . . in the philosophical metabolism' which mediates the mentalities of Rawls and the tradition of German Idealism in which Habermas is steeped, but that is decisively rejected by Cohen (AS: 151-52).

I have suggested that Cohen's position represents a Platonic foundationalism. This does not mean that Cohen's is a comprehensive foundation, an all-embracing perfectionism. It is rather a logical foundationalism, concerned with the formal structure of rational arguments and the necessity of fundamental propositions in that structure (Vincent 2004: 3-7). In Rawls's terminology, a Platonic approach to political philosophy is based on a Cartesian methodology that seeks self-evident first principles as a basis for deductive reasoning. In Chapter 5, I will argue that Cohen is committed to a form of 'Platonic atomism' that defined analytic philosophy in the moment of its formation through Bertrand Russell's and Moore's rejection of the Hegelian idealism of T.H. Green and F.H. Bradley (Hylton 1990).³³

Platonic atomism holds that propositions are objective entities that are independent of our acts of judging. That there are conditions or presuppositions of our acts of judgement does not mean on this view that these are conditions or presuppositions of the proposition or of the object of judgement. As a result, when the facts change the status of fundamental principles remains unchanged. Cohen remarks, 'facts, alas, are not my field. Being a political philosopher of the Anglophone stamp, a dealer in arguments, facts, for me, belong to minor premises the truth of which it is not my business to evaluate' (Cohen 1994: 1). Moral and political philosophy are for him 'ahistorical disciplines which use abstract philosophical reflection to study the nature and truth of normative judgements' (SFE: 1). This suggests one interest that Cohen might have in refuting Korsgaard's constructivism. Korsgaard's problem of reflexivity arises for Cohen in the first-person standpoint as a relationship between a personal Marxist normative conviction and changing social circumstances. Cohen argues that at the level of first-person reflection we cannot believe that our convictions are true *because we have bound ourselves to them*. 'Anyone who believed in [socialist values] believed in them because she thought them inherently authoritative, and the withering away of the social forces that backed them cannot justify ceasing to think them authoritative' (Cohen 1997: 7).

Cohen combines a personal Marxist normative conviction with a personal philosophical conviction that in philosophy personal convictions are fundamental.³⁴ The

³³ Unlike Hegel, 'Plato and Descartes . . . insisted on comprehensibility without giving it [a] dialectical twist' (IYE: 62).

³⁴ To neglect the autobiographical reflections that Cohen makes central to his mode of argumentation on the assumption that they are irrelevant to Cohen's 'real' philosophical

paradigm of the injustice of capitalist exploitation is for him a moral datum (SFE: 145), the product of a political upbringing with a strong resemblance to a religious upbringing (IYE: 7). In addition, he remains committed to the analytic/synthetic distinction emblematic of analytic philosophy, acquired during a post-graduate education at Oxford. Cohen describes these as two paradoxes of conviction. What is taken to be paradoxical is that rationality would seem to require that reflecting on the contextual origins of our fundamental convictions should place in doubt our rational grounds for believing them. But at the reflexive level of the first-person standpoint we cannot believe this to be true. That we often lack grounds for belief that are superior to those that others have for their conflicting beliefs contradicts

an undeniable principle about reasons for belief – namely, that you lack good reason to believe p rather than a rival proposition q when you cannot justifiably believe that your grounds for believing p are better than another's are for believing q (IYE: 11).

However, one might argue that one lacks good reasons to believe something when those beliefs are not the product of procedures which one has good reasons for believing are conducive to the production of correct beliefs. Cohen can be faulted for failing to address why his particular philosophical course offers privileged insights (Callinicos 2001). Reasons for normative beliefs can be distinguished from reasons for empirical beliefs in being held not at the reflexive level of the first-person but rather at the reflexive level of the first-person plural from which we consider the ‘reasons we can share’ (Korsgaard 1996b). This relativization of belief in relation to the standpoint of others does not threaten to undermine our everyday beliefs. It rather distinguishes between beliefs *per se* and beliefs that may justifiably be acted upon for moral and political purposes. Cohen recognises the depth of his methodological disagreement with Rawls: it is about fact and value, the relation of theory and practice and the status of philosophy itself (RJE: 4). He defends an ‘Oxford’ approach in which ‘we determine the principles that we are willing to endorse through an investigation of our individual normative judgments on particular cases’ against a ‘Harvard’ approach that assumes philosophy can move further away from prephilosophical judgement (RJE: 3-4). But he does not remark upon the converse of this, which is that the ‘Oxford’ approach assumes as against a ‘Harvard’ approach that the systematizing philosophical methodology through which principles are established *can* move further away from everyday judgements.

arguments would be to fail to subject to scrutiny highly contestable philosophical assumptions about the relationship between philosophy and biography.

To the extent that Rawls and Habermas both oppose the methodological search for foundational logical first principles they are non-foundationalists, but since this is not a debunking anti-foundationalist relativism it is best described as an immanent foundationalism, the search for a self-justifying foundation that pulls itself up by its own bootstraps via a virtuous circle of reasoning (Vincent 2004: 4). From the perspective of traditional analytic philosophy, this seems an obfuscating refusal to take a stand, as absurd as Baron Münchhausen's claim to have pulled himself out of a swamp by his hair. However, such a bootstrapping position becomes more attractive when one considers the problems that beset foundationalism. What Hans Albert calls Münchhausen's trilemma is so-called because in pursuing a foundationalist strategy the three justificatory alternatives with which one is faced appear equally unacceptable. '[I]f one demands a justification for *everything*, one must also demand a justification for the knowledge to which one has referred back the views initially requiring foundation'; one must therefore choose between:

1. an *infinite regress*, which seems to arise from the necessity to go further and further back in the search for foundations, and which, since it is in practice impossible, afford no secure basis;
2. a *logical circle* in the deduction as in the grounding process statements are used that were already characterized as in need of grounding, likewise leading to no secure basis; and, finally,
3. the *breaking-off of the process* at a particular point, which admittedly, can always be done in principle, but involves an arbitrary suspension of the principle of sufficient justification (Albert 1985: 18).

Albert's critical rationalism follows Karl Popper in avoiding the trilemma by giving up the principle of sufficient justification, so-called after Leibniz's principle of sufficient reason, which regards an adequate foundation as having the certainty of an Archimedean point. All justification, normative and scientific, rests not on a solid bedrock but in a Popperian sense on piles driven down into a swamp, sufficient for now to bear the weight of one's theory (Popper 2002 [1959]: 94). The similarities between the principle of sufficient justification and Cohen's 'Principle' are striking, and Cohen, like Albert, argues that we have to break off the process of justification, although whereas in a Popperian view investigation can be re-opened, for Cohen normative justification terminates in personal convictions.

Habermas's critique of Albert will assist in highlighting the deficiencies of Cohen's view. Conceiving of the interrelation of rational convictions and the context in which those convictions were acquired as a problem ideally to be overcome blocks reflection on the social presuppositions of knowledge and reasoning, the background that gives commitment to reason its meaning. Habermas argues against Albert that the

circle resulting from the internal relations between the subject and the object of reflection has to be treated dialectically, dialectical theories incorporating reflexively ‘the fact that they themselves remain a moment of the objective context which, in their turn, they subject to analysis’ (Habermas 1976 [1963]: 134). Habermas later treats this Hegelian dialectical insight with greater reference to Hans-Georg Gadamer’s hermeneutics (Holub 1991), the key point being that while Albert holds that all views are accountable to criticism, he does not demand that all views, including the critical attitude itself, be rationally grounded. This becomes untenable when we recognise that the pre-understandings that are the conditions of critical testing can themselves be criticised. Albert assumes that the problem of foundation (*Begründungsproblem*) can be resolved through a critical rationalism that abstains from the problem of self-foundation (*Selbstbegründung*). But if one breaks with the confines of a deductive paradigm of justification, argumentation becomes reflexive. The principles on the basis of which criticism proceeds can themselves be subjected to criticism. Criticism cannot be defined because the standards of rationality can only be explained within criticism itself. ‘This is the dimension of comprehensive rationality which, although incapable of final grounding [*Letztbegründung*], develops in a circle of reflexive self-justification’ (Habermas 1976 [1964] : 214).

The suggestion that Rawls shares such a view might seem far-fetched. But Rawls’s understanding of justification through reflective equilibrium is part of a tradition of American holistic pragmatism (White 2002) influenced by a naturalized Hegelianism. Reflective equilibrium is a circular approach to justification, influenced by the pragmatism of Nelson Goodman, Morton White and W.V.O. Quine, that responds to Münchhausen’s trilemma (Føllesdal 2005). Rawls’s understanding of the role of facts in justifying principles represents a Hegelian rejection of the dualisms of Kant’s transcendental philosophy, between the analytic and the synthetic, a prior vs. a posteriori and pure vs. empirical uses of reason (Freeman 2007a: 26). It is Hegelian in being Quinean, for not for nothing did Dreben describe Quine as the ‘Hegel of contemporary philosophy’ (Franks 2007: 66 n.1), compare Rawls’s methodological holism with Hegel’s, and say of *TJ* that it ‘reads like it was translated from the original German’ (Freeman 2007a: 28).

The lesson of reflective equilibrium is, as Kai Nielsen puts it, that

we cannot escape starting from tradition and from some consensus. In this fundamental sense we unavoidably start from morality as *Sittlichkeit* and refer back, however far we go in a reformist direction, to that *Sittlichkeit* ... Here, in spite of what otherwise are enormous differences, Rawls and a

hermeneuticist such as Gadamer have a common point of departure (Nielsen 1987: 145).³⁵

Like in Otto Neurath's parable of the mariner at sea, we cannot replace all of the planks of the boat at the same time in accordance with the principle of sufficient justification. What we can do is seek to provide a rational justification of the need for, and acceptability of, the sea of social practices upon which we are afloat.

Nielsen interprets reflective equilibrium in terms of an emancipatory perspective. This recalls Habermas's early theorisation of an emancipatory interest in *Knowledge and Human Interests* (KHI). However, it is important to take account from the start of Habermas's paradigm shift from the philosophy of consciousness to the philosophy of language. In a self-critique in a postscript, Habermas argues that

the traditional use of the term reflexion, which goes back to German Idealism, covers (and confuses) two things: on the one hand, it denotes the reflexion upon the conditions of potential abilities of a knowing, speaking and acting subject as such; on the other hand, it denotes the reflexion upon unconsciously produced constraints to which a determinate subject . . . succumbs in its process of self-formation (KHI: 377).

Habermas's rejection of the dialectical approach of self-criticism in ideology critique in favour of a Kantian conception of rational reconstruction inaugurated a transitional stage in which he worked on developing a new paradigm for critical theory, influenced in taking the linguistic turn by Apel's transcendental pragmatics of language (Apel 1975). However, while Habermas appeals to Apel's answer to Münchhausen's trilemma in his discourse ethics (MCCA: 79), the definitive moment in the development of his mature position is the social theory of *The Theory of Communicative Action* (TCA 1, TCA 2). While Apel believes that a pragmatics of language keeps alive the project of fundamental grounding, Habermas rejects from the start Apel's foundationalist search for ultimate justification, arguing that philosophy is not independent of social theory. Habermas is therefore best seen as seeking to reconcile Albert and Apel, reconceiving of the traditional ambitions of idealist philosophy in the pragmatic spirit of critical rationalism.

³⁵ On Rawls and hermeneutics, see Josefson (1997); for criticism see Vergés-Gifra (2006).

2.4 Hegel and the Philosophical Problem of Modernity

While a Hegelian ‘having one’s cake and eating it’ combination of description and prescription is liable to arouse suspicion, it has well-defined motivations in seeking a postfoundationalist answer to Kant’s conception of self-legislation. The non-traditional or post-Kantian view of Hegel with which I am concerned downplays Hegel’s metaphysical teleology of *Geist* coming to know itself (Taylor 1975), seeing Hegel as taking forward Kant’s project of understanding how the authority of reason can rest on nothing other than reason itself.³⁶ It also recognises that despite Hegel’s own criticisms of Rousseau (Ripstein 1994), ‘Hegel is fundamentally indebted to Rousseau in his social conception of freedom, inheriting a tradition that begins with Rousseau’ (Pippin 1997a: 93).

It is uncontroversial to situate Habermas’s position in relation to this framework. Pippin (1999) follows Habermas when he describes how with Hegel the problem of self-grounding becomes a social and historical problem: the philosophical problem of modernity. While it may be thought alien to Rawls’s view, I follow Sibyl Schwarzenbach in believing that ‘Rawls’s theory may appropriately be labelled “Hegelian” . . . once we have granted . . . the possible separation of political theory from a full blown metaphysics’ (Schwarzenbach 1991: 542). Rawls follows proponents of the ‘new Hegel’ (Gillespie 1997), believing ‘most of [Hegel’s] moral and political philosophy can stand on its own’ independent of Hegel’s metaphysics (LHMP: 330).

This interpretation must be distinguished from the most well-known ‘Hegelian’ interpretation of Rawls, that of Rorty (1991a), outlined in opposition to Michael Sandel’s critique of Rawls’s Kantian reliance on an ‘unencumbered self’. Rorty’s position represents little advance over the terms that dominated the liberal-communitarian debate, a caricatured opposition between ‘Kantian’ ahistorical impartialist liberals and ‘Hegelian’ contextualist communitarians. Rorty presents a three-cornered debate, subdividing Hegelian critics of Kantianism into communitarians like Alasdair MacIntyre who oppose the institutions of modern liberal democratic societies and pragmatists (postmodernist bourgeois liberals) such as John Dewey who seek to defend these institutions independently from their traditional ‘Kantian’ justification. But,

³⁶ Following Hartmann (1972), this is sometimes referred to, perhaps infelicitously, as a non-metaphysical Hegel. For overviews see Redding (2008); Lumsden (2008). There are realist critiques of constructivist interpretations of Hegel as much as of Kant; see Stern (2007). I will not consider whether this line of interpretation either distorts Hegel by reading him through the lens of Kant, or Kant by reading him through the lens of Hegel, or both. However, even sympathetic commentators such as Pinkard (1990) have suggested that Pippin presents an overly Kantian view of Hegel.

as Rorty recognises, these two positions ‘take over Hegel’s criticism of Kant’s conception of moral agency, while either naturalizing or juking the rest of Hegel’ (Rorty 1991b: 197-98). It is both more accurate and more attractive a view to see Hegel as already ‘beyond liberalism and communitarianism’ (Williams 2001).

Rorty argues that the problem of self-grounding is not a problem but rather ‘scratching where it does not itch’, following Hans Blumenberg in arguing that the modern political project of self-assertion can be separated from the modern epistemological project of self-foundation or self-grounding (see also Mouffe 2005: 41).

According to Blumenberg, with Descartes

a claim was made to the absolute beginning of the modern age, the thesis of its independence from the outcome of the Middle Ages, which the Enlightenment was to adopt as part of its own self-consciousness. The exigency of self-assertion became the sovereignty of self-foundation, which exposes itself to the risk of being unmasked by the discoveries of historicism, in which beginnings were to be reduced to dependences (Blumenberg 1985: 184).

But it is both unconvincing and partial to identify the demand of absolute freedom with a Cartesian quest for first principles (Pippin 1997b). Examining the social turn inaugurated in Hegel’s understanding of the Kantian problem of self-legislation will prepare the way for an examination of Hegel’s historical view of the problem of self-grounding as the problem of modernity.

Kant argued that the authority of pure reason rests on nothing but pure reason itself. For Kant’s successors, then, reason’s being ‘absolutely supreme’ meant its possession of a ‘self-authorizing’ authority. However, in his practical philosophy, Kant came to understand the unconditional claim of the moral law as depending upon a ‘fact of reason’. On a reading of Kant’s commitment to self-legislation that post-Kantian Hegelians share with Kantian constructivists like Korsgaard, one rejected by traditional Kantians like Ameriks, this leads to a ‘Kantian paradox’:

The paradox arises from Kant’s demand that, if we are to impose a principle (a maxim, the moral law) on ourselves, then presumably we must have a *reason* to do so; but, if there was an antecedent reason to adopt that principle, then that reason would not itself be self-imposed; yet for it to be binding on us, it had to be (or at least had to be ‘regarded’ to be, as Kant ambiguously stated) self-imposed (Pinkard 2002: 59).

As Pippin puts it, ‘If person, or subject, or agent or spirit are normative terms, ideals we come to hold ourselves to, the result of instituting and sustaining rules and proprieties, then we can be subject to such a norm only by in some sense subjecting ourselves to it’

(Pippin 2008: 18). Hegel argues for the inseparability of subjective, individual good and its internal reasons and objective, common good and its external conditions. On his social and historical conception of self-legislation, objective rationality is the ‘achievement of a kind of success in the institution-bound practices of justifying ourselves to each other, “the game of giving and asking for reasons”’ (Pippin 2008: 32, citing Robert Brandom). Pippin admits that this attempt to reconcile dualisms ‘can sound like a “have your cake and eat it too” inclusiveness’ (2008: 33). But it can hope to allay the misgivings of Cohen: it ‘does not suggest: “I am only bound because I bound myself, so I hereby unbind myself”’ (2008: 74).

We saw with Rousseau the importance of the distinction between the natural and the social. In pursuing a post-Kantian problem of self-legislation, Hegel can be understood as taking forward a Rousseauian conception of social freedom (Neuhouser 2000), outlining a developmental, historical story which seeks to answer the question of what normative forms are consistent with the non-natural, rationally self-determining status of human subjects. ‘Like Kantian constructivists, the answer must have to do with those norms that realize and embody the condition of their very possibility: that they exist and are maintained only as legislated, only in so far as “spirit is a result of itself”’ (Pippin 2008: 64). Once, in a move that Korsgaard herself seems implicitly to have made, we regard subjectivity as socially and historically formed, then practical identity will require a developmental rather than a deductive basis. If this is the case then

Hegel may prove to be more Kantian than Kant . . . [T]he obvious Kantian thing to say here is that there is no particular reason to grant our historical location in late modernity any normative authority unless that form of life can be itself reflectively endorsed (Pippin 2008: 91).

For Hegel, then, the modernity problem is autonomy, and the idea of self-determination or self-grounding is the principle of modernity (Pippin 1999: 12-13). This principle requires a criterion of self-determination, ‘a criterion we cannot simply be said to share by being human, or to find inscribed in Platonic heaven’ (1999: 14). Pippin follows Habermas’s claim that ‘Hegel was the first to raise to the level of a philosophical problem the process of detaching modernity from the suggestion of norms lying outside of itself in the past’ (PDM: 16). The ‘problem of self-grounding’ becomes the ‘problem of grounding modernity out of itself’ (PDM: 8).

Modernity can and will no longer borrow the criteria by which it takes its orientation from the models supplied by another epoch; *it has to create its normativity out of itself*. Modernity sees itself cast back upon itself without any possibility of escape (PDM: 7).

The source of the need for philosophy for Hegel is the need for self-reassurance caused by the anxiety created by ‘the fact that a modernity without models had to stabilize itself on the basis of the diremptions . . . it had wrought’ (PDM: 16). The fundamental question is, then: ‘How can an intrinsic ideal form be constructed from the spirit of modernity, one that neither just imitates the historical forms of modernity nor is imposed upon them from the outside?’ (PDM: 20).

2.5 Rawls’s Hegelianized Kantianism

The claim that the projects of Rawls and Habermas are best interpreted in relation to the problem of self-grounding is not without support in the literature. Larmore recognises that Rawls and Habermas share the aim of developing a self-standing or autonomous conception of the guiding principles of modern democracy (Larmore 2008: 153). But while Larmore criticises Habermas’s attempt to do so, his remarks on Rawls are offered in the spirit of clarification. Larmore detects a tension between the principles of justice that develop into Rawls’s political liberalism and the Kantian conception of autonomy on the basis of which they were defended in *TJ*. The claim that in responding to the problem of self-grounding Rawls offers a Hegelianized Kantianism therefore requires defence. After doing so, I discuss Habermas’s critique of Hegel, and suggest how similar criticisms might apply to Rawls.

Rawls’s ‘Kantian interpretation’ of justice as fairness has been subjected to severe criticism.³⁷ However, little attention has been paid to what Rawls means by explicitly presenting an *interpretation* of Kant. For Rawls, amending or adding to Kant can be ‘Kantian in spirit, provided that . . . it doesn’t compromise the essential elements of his doctrine’ (LHMP: 174 n.4): ‘the adjective “Kantian” expresses analogy and not identity’ (CP: 304). Subscribing to a Kantian *ethos* – an ethos of autonomy – does not mean endorsing Kant’s own position (Tampio 2007). The distinctive character of Kant’s theory is not lost by abandoning its deep metaphysical dualisms. Indeed, Rawls aligns himself not simply with Kant but with a tradition of the liberalism of freedom comprising Kant, Hegel and, less straightforwardly, John Stuart Mill.³⁸ This contrasts

³⁷ See Johnson (1974, 1977); Levine (1974); Mason (1976). Cf. Darwall (1976).

³⁸ I follow Freeman in emphasising the importance of Rawls’s engagement with the canon. But in identifying Rawls’s idea of a well-ordered society with Rousseau, Kant and ‘the German Idealists, including Marx’ (Freeman 2007a: 461), Hegel is conspicuous by his absence, and Freeman only grudgingly acknowledges Rawls’s inclusion of Hegel in the tradition of the liberalism of freedom (Freeman 2007b: 103). While it may be true that ‘Hegel had little *direct* influence on Rawls’s *initial* working out of the main ideas of *A Theory of Justice*’ (Freeman 2007a: 25, emphases added), it is plausible to claim this influence was indirect through the work

with the tradition of the liberalism of the good to which classical utilitarians subscribe. Putting the accent on Rawls's methodological approach justifies adding Rousseau to the start and dropping Mill from the end of this genealogy.³⁹ I will argue Rawls's project is best understood in terms of the constructive appropriation of a tradition of thought that runs through Rousseau, Kant and Hegel – what, following Rawls's identification of the 'BES line' (LHPP: 375) of classical utilitarianism (Bentham, Edgeworth and Sidgwick), we can call the RKH line.

The Rousseauian dimensions of Rawls's project are a key theme of the work of Joshua Cohen, and have been noted by others (Dent 2005: 227-29; Neal 1997: 51-71; Manin 1987). Rawls's affinity with Hegel has been recognised by Hegel scholars (Neuhouser 2000; Hardimon 1994), who in turn influenced Rawls's interpretation of Hegel (LHMP: 331 n.4); by commentators on Rawls (Audard 2007; Kukathas 1990); and by those with an interest in both (Schwarzenbach 1991; Sensat 2003, 2007).⁴⁰ As has the way Rawls draws on ideas that unite Rousseau and Hegel (Neuhouser 2000; Laden 2001, 2007), the interpretations of Nicholas Dent (1988) and Neuhouser (1993), which emphasise how the Hegelian idea of recognition is prefigured in Rousseau, influencing Rawls's interpretation of Rousseau. However, so far no interpretation has placed Rawls within the framework of the post-Kantian Hegel, and therefore the full RKH line. This is despite Schwarzenbach's conclusion that *TJ* 'stands squarely in the tradition of the German *Rechtslehre*' (Schwarzenbach 1991: 550), and John Chapman's observation that:

In the perspective of western moral and political philosophy . . . Rawls's theory of justice is best seen as the culmination of the effort, begun by Kant and Hegel, and carried forward by T.H. Green, Bernard Bosanquet, and the other British and American idealists, to adapt Rousseau's theory of the general will to the modern state (Chapman 1975: 588).⁴¹

O'Neill's observation that Rawls's constructivism is 'more Rousseauian than Kantian, more civic than cosmopolitan' (O'Neill 2003a: 353) should therefore come as

of Bradley and Dewey (Schwarzenbach 1991: 565 n.4). Freeman (2002) departs from Rawls's description of Hegel as a 'moderately progressive reform-minded liberal' (LHMP: 330, 349).

³⁹ Rousseau lacks a conception of individual subjectivity that is essential to Mill as well as Kant and Hegel (JF: 143).

⁴⁰ See also Benson (1994); Schwarzenbach (1994); Ivison (2002); Rasmussen (2004); Ege (2008); Beiner (2009). Criticisms of Rawls's ahistorical abstraction are sometimes even combined with recognition of the Hegelian influence (Plant 1991; Geuss 2008). See also the recognition of similarity in Hegelian critiques of Rawls in Hoy (1981); Peddle (2000, 2004); Buchwalter (2001); Houlgate (2001).

⁴¹ '[F]ew lines of affiliation are better established in the history of philosophy than that between Rousseau's declaration that liberty is the quality of man and the philosophy of Right as it developed from Kant to Hegel' (Bosanquet 2001 [1899]: 151).

no surprise. Rawls focuses on Kant's conception of autonomy and, in defending Kant against the criticisms of Sidgwick and idealists, argues that '[i]t is a mistake . . . to emphasize the place of generality and universality in Kant's ethics' (TJ: 251/221). While O'Neill represents Rousseau and Kant as polar opposites, Rawls appeals to Ernst Cassirer's Rousseauian interpretation of Kant, which stands as an important precursor to Kantian constructivism in arguing that 'the social bond must bind together freely acting persons, not dead things . . . it cannot be something imposed upon the wills of those persons from without; they must constitute and create it themselves' (Cassirer 1945: 31).⁴² Otfried Höffe describes Rawls's purported Kantian interpretation as Rawls's 'annoying claim', arguing Rawls's neglect of Kant's concern with positive law leads to a view of political community closer to Kant's view of religious community (Höffe 1984). But again we find Rawls reading Kant through the prism of Rousseau. Rawls detranscendentalizes, secularises and democratises Kant's Kingdom of Ends; it becomes for Rawls the 'realm of ends [which] is a secular ideal' (LHMP: 312). And Rawls draws on Kant's *Religion Within the Limits of Reason Alone* in order to present Kant as subscribing to an Augustinian moral psychology, rather than the Manichean metaphysical dualism suggested by the *Groundwork*, in a manner that allows him to suggest that Kant shares Rousseau's conception of both the problem and the potential cultivation of *amour propre* in the formation of moral community (LHMP: 291-94).

In accordance with his lexical ordering of his two principles of justice, Rawls's primary commitment is to a conception of equal freedom, or political autonomy, and only derivatively to distributive equality, following Rousseau in seeing distributive equality as constitutive of equal freedom (TJ: 511/447; LHPP: 244-48). Rawls is indeed best situated within a republican as opposed to a liberal tradition (Laden 2006; de Francisco 2006; Audard 2007).⁴³ The fact that principles of justice apply to the basic structure of society and specify a public point of view that is mutually recognized by all citizens means that Rawls's strategy of justification departs from the Hobbesian strand of the liberal tradition and adopts instead an approach also found in Rousseau's *Social Contract* and Hegel's *Philosophy of Right* (CP: 426 n.10). Following Dewey in seeking

⁴² Rawls is in line with Cassirer (1945: 59) in seeing Kant as seeking to give Rousseau's ideas a systematic foundation. Rawls cites Cassirer (1945) in TJ: 264/omitted and cites Cassirer (1954) in commenting: 'As so often, Kant is the best interpreter of Rousseau' (LHPP: 200).

⁴³ Rawls's identification as a liberal owes more to commentators such as Barry (1973), Wolff (1977) and Dworkin than to Rawls himself. The only entry for liberalism or its cognates in the index of *TJ* is to liberal equality, which Rawls rejects in favour of democratic equality; see Holbo (2008). Rawls says that while it is inconsistent with civic humanism, justice as fairness as a form of political liberalism is 'perfectly consistent with classical republicanism' (JF: 144; see also PL: 205). However, given that in contemporary political theory, republicanism has become associated with the neo-Roman republicanism of Pettit (1997) and Skinner (1998), it is more straightforward to focus on Rawls's appropriation of the RKH line independently of whether this is republican, liberal or a combination of the two.

to overcome Kantian dualisms (CP: 304), Rawls's detranscendentalized procedural rendering of Kant's categorical imperative issues in a choice of principles that are realized, in Hegelian fashion, in the basic structure of society.

Driven by a Rousseauian 'democratic impulse', Rawls's constructivism 'aims to generate the justification of justice, not to impose it from the start'.

The genius of Rawls is to switch the onus on to the justification process itself so that the moral doctrine of equal respect for persons can be adopted as a *result* for political purposes, not as a starting point . . . Rawls' political philosophy is new in that [he] places respect for autonomy *practically*, not *theoretically*, at the heart of his doctrine in the method of justification itself (Audard 2007: 7-8).

According to Larmore, though, Rawls never made clear the 'foundational' role of respect in his conception of justice (Larmore 2003: 383). Larmore recognises that for Rawls the notion of respect, or of the inherent worth of persons, is not a suitable basis for arriving at principles of justice (TJ: 586/513), but argues Rawls 'fails to grasp the real nature and import of the moral assumptions that drive his project' (Larmore 2008: 143).⁴⁴ However, Rawls's description of a political conception of justice as 'freestanding' courts misunderstanding. The designation 'freestanding' emphasises the negative point that a political conception seeks independence from comprehensive doctrines. But more perspicuous is Rawls's later description of a political conception as *self-standing*, grounding itself from the values of the domain of the political (CP: 611, 617). Larmore's emphasis on the need for a 'foundation whose authority must be understood as antecedent to the collective will of the citizens' (Larmore 2008: 150) contrasts with Rawls's claim that for constructivism, 'the union of practical reason and conceptions of society and person is complete and independent. For moral purposes, this union needs no grounding, and the principles of practical reason rely on nothing prior to it' (LHMP: 241).

There are therefore strong grounds for claiming that, working within a Rousseauian paradigm, Rawls's project of Hegelianized Kantianism is driven by the problem of self-grounding. But if this is the case, we may ask whether Rawls overcomes the aporias of self-referential reflective consciousness that lead Habermas to advocate a paradigm shift to the philosophy of language. Does Rawls, like Hegel, solve the problem of self-grounding *too well* (PDM: 42)? According to Habermas:

⁴⁴ For Hampton (1980: 338) too Rawls is 'more fully Kantian than he realizes'. See also Hampton (1994).

Hegel has ultimately to deny to the self-understanding of modernity the possibility of a critique of modernity. The critique of a subjectivity puffed up into an absolute power ironically turns into a reproach of the philosopher against the limitations of subjects who have not yet understood either him or the course of history (PDM: 22).

Denying the possibility of such a critique is precisely to deny to modernity the impulse of subjectivity that itself set modernity in motion: ‘absolute spirit neutralizes the conditions under which modernity attained a consciousness of itself’ (PDM: 43). This is not of course to suggest that Rawls is an absolute idealist. More pertinently, and recalling his criticism of Rousseau, Habermas criticises Hegel’s ‘emphatic institutionalism’, as well as his aspiration that one’s own standpoint is ‘to be brought to reflective awareness within the horizon of history as a whole’ (PDM: 6).

Habermas takes Hegel to task for assuming what has to be proven, that reason has an absolute power of unification, and therefore failing to proceed in a rigorously immanent fashion. Pippin (1997c) rejects this criticism, believing that for Hegel, the reconciliation promised ideally by modernity is not something that can be achieved once and for all (Pippin 1999: 75). I cannot enter this interpretative dispute here. Three points will suffice in the present context. First, it seems preferable to follow Habermas in distinguishing between the historical Hegel and the view propounded by contemporary interpretations influenced by Hegelian trends in post-analytic philosophy (Gordon 2005). Rawls’s synthetic appropriation of the canon can be criticised in similar terms to Pippin’s. Second, the problem with which Habermas is concerned appears more acutely when one adopts the perspective of a critical social theory, and the potential self-subverting tendencies of the dialectic of Enlightenment (Horkheimer and Adorno 1997 [1944]), rather than confining oneself to textual interpretation. Finally, it is worth noting the final thinker that Habermas discusses as developing the philosophical discourse of modernity: the realist sociologist and systems theorist Niklas Luhmann, who I discuss in Chapter 9. From this post-Weberian perspective, the ideal of being ‘at home’ in the institutions of the modern world takes on both an unrealistic and unattractive cast (Baynes 1995; cf. Pippin 1995).

2.6 Marx and Socialism, Utopian and Scientific

Marx, and Marx’s relationship to Hegel, is important to my concern with the rational agency tradition in relation to two stylised and interrelated sets of contrasts. The first is between Marx’s materialism and Hegel’s idealism, while the second is between Marx’s revolutionary stance and Hegel’s reformism. Of course, interpretations of Marx

regarding the latter point vary – Rawls (TJ: 281 n.19/523 n.4, 249 n.18/460 n.4) contrasting the revolutionary interpretation of Tucker (1969) with the Hegelian interpretation of Avineri (1969) – but Cohen endorses the revolutionary view, leading him to reject the Hegelian problem of self-grounding.

Both Rawls and Habermas endorse aspects of Marx's materialism. Rawls, following Cohen in interpreting Marx as holding a theory of justice (LHPP: 342-46), sees Marx as a critic of liberalism, against which justice as fairness must defend itself. Rawls's rejection of welfare state capitalism in favour of the idea of property-owning democracy (Krouse 1988) is central to his response to Marxism (LHPP: 320-21).⁴⁵ Justice as fairness is not, as Cohen contends, 'neutral . . . between capitalism and socialism' (RJE: 163); rather, historical circumstances and traditions determine the choice between property-owning democracy and liberal (democratic) socialism (JF: 139). Rawls emphasises the importance of maintaining the fair value of political liberties, as opposed to their being merely formal,⁴⁶ and the role of the institutions of a property-owning democracy in maintaining background justice. It is an irony that does not escape Cohen's notice (IYE: 183 n.2) that while for theorists like Barry, Rawls's focus on the basic structure of society 'represents the coming of age of liberal political philosophy' by integrating into an individualistic tradition the legacy of the structural views of Marx and Weber (Barry 1995a: 214; see also Gutmann 1989), this focus is targeted in Cohen's ostensibly Marxian critique of Rawls.⁴⁷

Rawls's brand of liberalism is a more democratic liberalism (J. Cohen 1994a) partly because it is a more materialist liberalism. Habermas's trajectory moves in the opposite direction. In seeking a more democratic materialism, Habermas draws on the radical-democratic Young Hegelianism of American pragmatism. If analytical Marxism is the project of updating Marxism using current methods of analytical philosophy and positivist social science, then Habermas as much as Cohen is an analytical Marxist (Heath 2009). However, their respective conceptions of historical materialism and the rational socialist society are at opposite ends of a spectrum (Smith 1984). For Cohen, socialism is the inevitable, because optimal, development of productive forces, while for Habermas democratic socialism is the possible outcome of moral learning processes (J.

⁴⁵ See TJ: 274/242; JF: 176-79. This is an institutional corollary of Rawls's methodological commitment to a liberalism of freedom over liberalisms of happiness (Freeman 2007b: 108). On property-owning democracy, see also Jackson (2005); Ron (2008); Martin O'Neill (2009). It has a Rousseauian pedigree (Rousseau 1997 [1755]: 147) and follows Cassirer (1954: 60). The idea of embedding the economy in surrounding institutions of public regulation is reminiscent of Hegel (LHMP: 357).

⁴⁶ See TJ: 221-228/194-200 and the reemphasis in response to the Marxian criticism of Daniels (1989) in PL: 324-331.

⁴⁷ Rawls (TJ: 259/229) reciprocally acknowledges Barry's (1965: 75-79) development of Marxist insights regarding the role of the social system in shaping wants and aspirations.

Cohen 1989a).⁴⁸ While for Cohen technological development and instrumental rationality drive human history, Habermas places changes in moral-practical consciousness in the separate domain of communicative rationality at the centre of historical development. If for Cohen analytical Marxism meant rendering Marxism compatible with the assumption of methodological individualism in neoclassical economics and game theory, for Habermas it meant engaging with the systems theory functionalism of Luhmann.

Here the issue of materialism joins with the second issue, the revolutionary character of Marxism. In the tradition of the Frankfurt School, Habermas was, and still is in some ways, a Hegelian Marxist, endorsing Marx's materialism against Hegel's idealism while rejecting Marx's singular focus on the revolutionary potential of social labour. For Habermas, Marx's practice of critique goes beyond his theory of historical materialism, serving as the inspiration for ideology critique that is Marxian in inspiration but employs Hegelian dialectical methods. Habermas was, and remains, a radical reformist. He sees Marx as succumbing to the same totality thinking as Hegel, translating the Hegelian conception of the self-realization of spirit into the domain of social labour in his conception of the revolutionary process that overcomes alienation (PDM: 342). Habermas is drawn to Hegel's early Jena philosophy, from which he derives a fundamental distinction between labour and linguistic interaction (TP: 142-70), and which provides a basis for reconstructing historical materialism in a more idealist fashion as a moral learning process resulting from the development of normative structures of communicative action (CES). In Habermas's mature theory, this reconstruction of historical materialism finally gives way to a more Weberian view according to which political, legal and cultural phenomena all exert a causal influence over the development of the productive forces (TCA 1, TCA 2). Under the changed premises of the paradigm of communicative rationality, 'there is no equivalent for the philosophy of the subject's model of self-influence in general and for the Hegelian-Marxist understanding of revolutionary action in particular' (PDM: 361). Marx's position must be defended on the basis of an alternative theoretical paradigm (Rockmore 1987).

Cohen, on the other hand, remains wedded to a revolutionary paradigm because for him it is a Marxian ideal of equality that is fundamental to Marxism, not historical materialism as a method (KMTH: xxii). Historical materialism is for Cohen a descriptive theory to be defended using the best tools available, and theoretical scepticism about historical materialism leaves a socialist vision unaltered (HLF: 132). Cohen initially

⁴⁸ See J. Cohen (1982) for a critique of Cohen's thesis that material development is independent of control by social structures. Cf. HLF: 83-109.

defended an ‘old-fashioned’ technologically determinist historical materialism (KMTH: x), but when its tenability was undermined – by the decline of the working class (IYE: 108) and the discrediting of the idea of guaranteed future material abundance – Cohen switched his focus decisively to normative theory (SFE: 6). This way was left open because Cohen always rejected the view that Marx did not condemn capitalism for being unjust (Cohen 1983) or thought principles of justice were irrelevant to the socialist movement (SFE: 3). The fact that for Cohen moral and political philosophy are ahistorical disciplines which inquire into the truth of normative judgments is the obverse of his view that historical materialism is a substantially value-free, and refuted, empirical theory (SFE: 1). Cohen advocates socialism making the return journey from science to utopianism, again writing recipes for the cookshops of the future (IYE: 77). ‘Marxism’s hard factual carapace is cracked, its soft underbelly of values is exposed’ (IYE: 103; SFE: 5-6). Cohen may have ‘moved from an economic point of view to a moral one, without ever occupying a political one’ (IYE: 3),⁴⁹ but justice still depends upon a revolution that cannot take place through the reform of existing social structures, if not a proletarian revolution then a revolution in the human soul.

The place of revolutionary change in Cohen’s thought leads him to reject the terms of the problem of self-grounding. Cohen interprets Marx’s idea, derived from Hegel, that the goal of philosophy is to realize itself in the world, in Platonic terms (IYE: 97-100). ‘For utopian socialism, theory is developed independently of the world, and practice is the attempt by the subject (thought of as counterposed to the world) to make the world conform to the demands of theory’ (IYE: 74). Marxism, even analytical Marxism, never eliminated the vestiges of Hegel in Marx, the ‘obstetric motif in the Marxist conception of revolution’, evident in Rosa Luxemburg’s contention that ‘History . . . has the fine habit of always producing along with any real social need the means to its satisfaction, along with the task simultaneously the solution’ (IYE: 71). Cohen contrasts the figure of the midwife who derives a solution from within with the engineer who applies a scheme of independent design (IYE: 61), but has no place for the figure of the social reformer or critic.

Cohen’s critique of Rawls begins from substantive rather than methodological issues, but Rawls’s substantive differences with Marxism are primarily traceable to methodological differences. In distinguishing the idea of the well-ordered society of

⁴⁹ There are two intermediate stages in this journey, first the move from an inclusive to a restricted historical materialism (HLF: 155-183), occasioned by the need to remedy Marxism’s inattention to ‘cultural phenomena of a spiritual kind’ (KMTH: xxviii), namely religion and nationalism (HLF: 132-155), and, second, Cohen’s initial engagement in normative philosophy in his critique of Robert Nozick’s libertarianism. The move to the fourth stage, the engagement with Rawls, results from Cohen prioritising distributive justice over the principle of self-ownership he takes Marxism to share with libertarianism.

justice as fairness from Marx's idea of a full communist society, Rawls argues that the latter is a society that has surpassed the circumstances in which questions of justice arise. By contrast, 'justice as fairness assumes that, given the general facts of the political sociology of democratic regimes . . . the principles and political virtues falling under justice of various kinds will always play a role in public political life' (LHPP: 321-22; see also TJ: 281, 523 n.4/249, 460 n.4). To the extent that Marx shares with Rousseau an ideal of free social relations, Rawls can accept this vision. It is the revolutionary and utopian aspects of Marx's position to which Rawls is opposed.⁵⁰ Rawls seeks to avoid both a Hegelian ideological accommodation to the status quo (JF: 4 n.4) and an ideological defence of sectarian values, an aspect of his concern with the problem of stability being to show that the maintenance of a just order is not dependent upon ideological support (CP: 326 n.4). A Young Hegelian commitment to radical reformism is a commitment that Rawls shares with Habermas.

Without taking a view on their exact importance within Marx's own work, it is possible to identify two blind spots caused by the strictures of analytical Marxism that persist in Cohen's later approach. Both relate to the concept of internal relations and both can be appreciated from the perspective of the historical antecedents to Marx's position. In a narrow sense, analytical Marxism is opposed to holism and in favour of methodological individualism: 'behaviours of individuals are always where the action is, in the final analysis' (KMTH: xxiv).⁵¹ But, approaching Marx from a Rousseauian perspective, Cohen can be accused of conceiving of interests atomistically and failing to appreciate that interests exist only as a result of internal relations between persons as integral parts of the community they constitute (Levine 1993: 33-34). Both Rawls and Habermas defend this Rousseauian conception of social relations in opposition to the Lockean individualist conception of self-ownership that Cohen takes to be central to Marxism as much as libertarianism (SFE: 159-64). Second, in a broad sense, analytical thinking is opposed to so-called dialectical thinking and in favour of the techniques of logical and linguistic analysis developed within analytic philosophy (KMTH: xvii-xviii). But, approaching Marx from a Hegelian standpoint, this is to ignore the dialectical nature of Marx's materialism, the internal relation between base and superstructure (Sayers 1990). I will argue in Part II that these two methodological commitments are highly consequential to Cohen's interpretation of Rawls.

⁵⁰ Levine (1977) employs Rousseau to argue against Rawls on this point, but a Rousseau read through the lens of Marx. Cf. Brudney (1997).

⁵¹ Note the allusion to IYE: 134-48.

3. Bringing Rawls's and Habermas's Self-grounding Projects into Dialogue

The purpose of my reconstruction of the development of the rational agency tradition in the previous chapter was to show how Cohen rejects, and Rawls and Habermas accept, the terms of the problem of self-grounding. In this chapter I will argue that the problem of self-grounding provides the most fruitful perspective from which to understand the projects of Rawls and Habermas and within which to interpret the 1995 debate between them in the *Journal of Philosophy*. The need for this reorientation will be highlighted by examining the limitations of existing frameworks of comparative assessment. This analysis brings together the results of the previous two chapters. A shared commitment to a Rousseauian paradigm of the relationship between theory and practice means that the justificatory problem of self-grounding is central to the projects of both Rawls and Habermas. However, there are clear and significant differences in the ways in which they seek to address this problem that derive from the different ways in which they seek to reconcile Kantian and Hegelian ideas, expressing competing commitments to realistic utopianism and utopian realism.

3.1 The Rawls-Habermas Debate

The standard view of the relationship between Rawls and Habermas sees Rawls as a Kantian liberal in the social contract tradition of Hobbes and Locke (CES: 184), and Habermas as relying upon a comprehensive philosophical doctrine in the mould of Hegelian grand theory. Indeed, it is this contrast which structures what Habermas calls their Kantian 'family quarrel' (HR 1: 50). More specifically, it lies behind what Rawls identifies as the two main differences between their respective positions: first, between their competing 'analytical devices of representation', his original position versus Habermas's ideal speech situation (PL: 381; see also LHPP: 19-20), and second a putative contrast between his political conception of justice and Habermas's comprehensive metaphysical doctrine. Thus, Rawls is concerned to defend himself against Habermas's accusation that the original position is used to deduce a set of pre-political moral rights, while in turn Rawls criticises Habermas for propounding a doctrine that is comprehensive in the Hegelian sense of seeking to provide an analysis of the presuppositions of all rational discourse.

However, closer attention to these two contrasts calls the standard view into question. The standpoint of the original position is a standpoint for choosing between

principles for giving an ideal form to the basic structure of our social world. It is, moreover, employed within a holistic process of justification through reflective equilibrium. The ideal discourse situation, on the other hand, is a means of representing the idealizations that are implicit within real procedures of argumentation. It does not have the ‘status of an *ideal* rooted in the universal presuppositions of argumentation and able to be approximately realized’, or against which we can identify ‘*deviations*’ (BFN: 322, 323), and these idealizations must be separately reconstructed in relation to truth, moral validity and legal validity. The relevant point of comparison is not therefore between the ideal discourse situation and the original position *simpliciter*, but between the reconstructive role of the ideal discourse situation and the constructivist use of the original position within a process of reflective equilibrium (Elster 1998: 5). Rawls recognises this to some extent when he says that reflective equilibrium resembles the ideal speech situation in being ‘a point at infinity we can never reach’ (PL: 385). Paraphrasing Marx, Rawls follows Hegel in descending from heaven to earth,⁵² while Habermas follows Marx in ascending from earth to heaven, clearly captured in the idea of ‘transcendence from within’ that Habermas has increasingly appealed to in preference to the ideal speech situation.

Second, regarding whether Habermas proposes a comprehensive doctrine, Joshua Cohen’s position is representative of the Rawlsian view that ‘political arguments should not be made to depend on, or presented as dependent on, a philosophical theory about the nature of reason’ as this does not ‘provide the common ground for equal citizens that is desirable in public arrangements in a democracy’ (J. Cohen 1999: 387). But it is only because J. Cohen, like Rawls, assumes that establishing common ground between free and equal citizens depends upon a philosophical theory of justice that he is concerned that this not be presented as depending upon a philosophical theory of general scope. These arguments for philosophical modesty are the corollary of a highly ambitious practical goal of providing substantive principles for the public arrangements of a democracy. In examining Habermas’s position, we will need to bear in mind that, as opposed to Rawls, Habermas has ‘never had any ambition of sketching out a normative political theory’, of designing ‘the basic norms of a “well-ordered” society on the drafting table’ (PF: 101).

Rawls and Habermas, and Rawlsians and Habermasians, have therefore tended not so much to engage with each other as with the ‘effective history’ (Gadamer 2003 [1960]: 267-74) of each other’s texts (Finlayson 2007: 147). In the Anglo-American world, Habermas is still regarded as a purveyor of grand theory with revolutionary Hegelian-

⁵² Compare Schaar (1974): ‘Starting as Marx might put it, from the heaven of the original position, Rawls is unable to descend to the earth of practice.’

Marxist ambitions (Skinner 2002). In Germany, on the other hand, Rawls's 'Kantian' theory has been interpreted through the prism of the liberal-communitarian debate (Müller 2002).⁵³ The mutually reinforcing influences of the work of Lawrence Kohlberg and Sandel has led to a view of Rawls as proposing a universalist and individualist moral theory (Finlayson 2007).

3.2 Kantian Family Dispute or Self-grounding Family Resemblance?

Three tendencies in the literature contribute to the standard view. They present (1) a moral interpretation of Rawls, (2) a misleading political interpretation of Habermas, or (3) a combination of the two. (1) The first approach sees Rawls and Habermas as offering rival Kantian approaches (Baynes 1992; Forst 2002). It is shaped by the context of the liberal-communitarian debate and the idea that Habermas offers a 'third way' between Rawls's impartialism and communitarian contextualism (O'Neill 1997; Outhwaite 1994). (2) The politicized version of Habermas has been articulated in relation to both Habermas's ideal speech situation and his discourse ethics. Early comparisons presented the ideal speech situation as an alternative device to the original position for deriving principles of justice or legitimacy (Pettit 1982; Benhabib 1982). And on the basis that Habermas describes moral principles as principles of justice, it has been argued that there is 'no difference between Rawls and Habermas' (McMahon 2002), overlooking the fact that while 'justice' for Habermas is a general moral term, for Rawls it refers to principles for the ordering of the basic structure of society.⁵⁴ A politicized view of Habermas is also offered by those who argue that Habermas's discourse ethics fails as a moral theory but can be read as 'a political theory in disguise' (Finlayson 2007: 146). (3) Appealing to T.M. Scanlon's contractualism as a bridge between Rawls and Habermas, the third approach seeks to synthesise a moralized Rawls and a politicized Habermas. The original position is presented in the mould of the ideal speech situation, or the moral principle of universalization (U), as a context-transcendent ideal that can be 'mirrored' in real processes of discourse modelled on a political interpretation of Habermas's discourse ethics. I criticise this view in Part IV.

Rather than interpreting the Rawls-Habermas debate as a Kantian family dispute, I will focus on the family resemblance in the way that the projects of Rawls and

⁵³ This is important to Habermas's 'early debate' with Rawls (Finlayson 2007). Habermas comments: 'Sandel has justly criticized Rawls for saddling the construct of an original position with the atomistic legacy of contract theory' (MCCA: 213 n.15). See also CES: 184, 198; MCCA: 66, 198.

⁵⁴ For criticism, see Rehg (2003); Finlayson (2007).

Habermas, both working within a Rousseauian paradigm, are driven by the problem of self-grounding. Their shared practical commitment to autonomy makes the methodological problem of self-grounding unavoidable. Such a view provides a new perspective on the developmental trajectories of their respective projects. Both Rawls and Habermas supposedly ‘placed a lot of early emphasis on ideals, and then retreated to more of an emphasis on process – when they found that those ideals were difficult to justify’ (O’Neill in Skinner 2002: 8-9).⁵⁵ However, from the perspective of the Rousseauian paradigm and its problem of self-grounding, I will argue that the overriding impression is of continuity rather than change. Second, it is not that with *PL* and *BFN* we have seen ‘a rapprochement from both sides’, with Habermas becoming more Kantian and Rawls more Hegelian (Müller 2002: 170). This confuses form and content. The content of Rawls’s conception, the principles of justice, have remained constant and Rawls’s justification of them has become more Hegelian. Habermas, on the other hand, begins with questions of justification, in the form of a conception of communicative reason, and becomes more Kantian when he becomes concerned with filling in the content of his conceptions of morality and law.

Employing Berlin’s distinction between two philosophical casts of mind – the hedgehog who knows one big thing and the fox who knows many small things – Rawls can be seen as a hedgehog and Habermas as a fox. Rawls’s realistically utopian political liberalism is a (Hegelian)political (Kantian)liberalism, or a Hegelianized Kantianism. Kantian constructivism is situated within the political context of our social world, conceived of in a Hegelian manner. Habermas’s utopian realist Kantian republicanism is a Kantian (Hegelian)republicanism, or a Kantianized Hegelianism. Habermas’s reconstructive moral and legal theories reconstruct Kantian principles implicit in a Hegelian context of communicative action. Their respective practical philosophies are only properly understood when seen in relation to their contrasting background metaphilosophies. By this I mean a conception of the reciprocal relation between the role of philosophy in modernity and a social ontology of the modern world. From this perspective, the projects of Rawls and Habermas do not confront one another on all fours. A comparative assessment will be more like distinguishing between evolving or degenerating research programmes (Lakatos 1978), or assessing how a tradition of moral enquiry deals with challenges (MacIntyre 1989).

In this comparison it is important not to elide the distinction between Rawls’s political-not-metaphysical liberalism and Habermas’s postmetaphysical philosophy.⁵⁶

⁵⁵ On Rawls, see Barry (1995b); on Habermas, see Scheuerman (1999).

⁵⁶ On Rawls, see Hampton (1989); on Habermas, Henrich (1999); on both, Williams (1999); Flikschuh (2000).

Rawls's position is incompatible with a Kantian transcendental metaphysics, for it equates philosophical metaphysical commitments with comprehensive religious doctrines, but not a Hegelian metaphysics of the social world. Political liberalism 'offers no specific metaphysical or epistemological doctrine beyond what is implied by the political conception itself' (PL: 10). One could even refer to this as a meta-metaphysics, an attempt to understand the social world in which metaphysical beliefs are formed in order to reconcile these beliefs (PL: 29 n.31; Boran 2005). From a Habermasian perspective, I will argue that the limitations of Rawls's approach, and its concern with reconciliation, derive from the metaphysical framework of the philosophy of the subject within which Rawls operates. Habermas understands metaphysics in terms of the reflexivity seen in German Idealism in which cognitive reason finds itself once again in the rationally structured world. But in opposing this framework, Habermas does not advocate a philosophical naturalism, appealing to a conception of transcendence immanent within language. Postmetaphysical thinking is better described as 'post-Idealist' thinking (Herbert Schnädelbach, quoted in Bowie 2003: 257 n.4). While Rawls therefore sounds anti-metaphysical but, I will argue, is metaphysical, Habermas sounds metaphysical in a pejorative sense but isn't. In concluding this chapter, I will briefly outline the metaphilosophies that frame the self-grounding projects of first Rawls and then Habermas (Figure 3).

3.3 Rawls: Political Liberalism as Hegelianized Kantianism

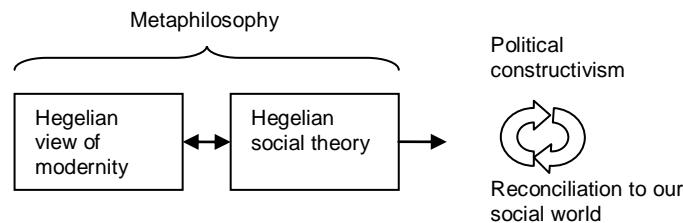
Rawls's reflections on moral philosophy begin from a contrast between classical philosophy's concern with the highest good and the concern with the prescriptions of right reason that marks modern moral philosophy (LHMP: 1-2). His own methodological conception of the role of philosophy in probing, in constructivist fashion, the limits of practicable political possibility in order to reconcile us to our social world is in the service of two practical aims of political philosophy within the public political culture of a modern constitutional democratic society. In Hegelian fashion, the first is the practical role of seeing whether despite political conflicts, reasoned bases for political agreement can be found, the work of abstraction being set in motion by deep political conflicts (PL: 44). Rawls, like Hegel, seeks to resolve the historical conflict between freedom and equality, the liberties of the moderns and the liberties of the ancients in Constant's terms (Hardimon 1994: 129 n.2). In pursuing how such a conception of justice can be justified in a pluralistic society, Rawls locates the historical origins of political liberalism in the Wars of Religion that followed the Reformation,

and which eventually led to a commitment to the principle of toleration. ‘As Hegel saw, pluralism made religious liberty possible, certainly not Luther’s and Calvin’s intention’ (PL: xxvi). Second, political philosophy plays the role of orientation, in a sense suggested by Kant’s use of the term, but understood in political terms as

how a people think of their political and social institutions as a whole, of themselves as citizens, and of their basic aims and purposes as a society with a history – a nation – as opposed to their aims and purposes as individuals, or as members of families and associations (LHPP: 10).

I argue in the next chapter that this social ontology is appropriately labelled Hegelian.

Rawls



Habermas

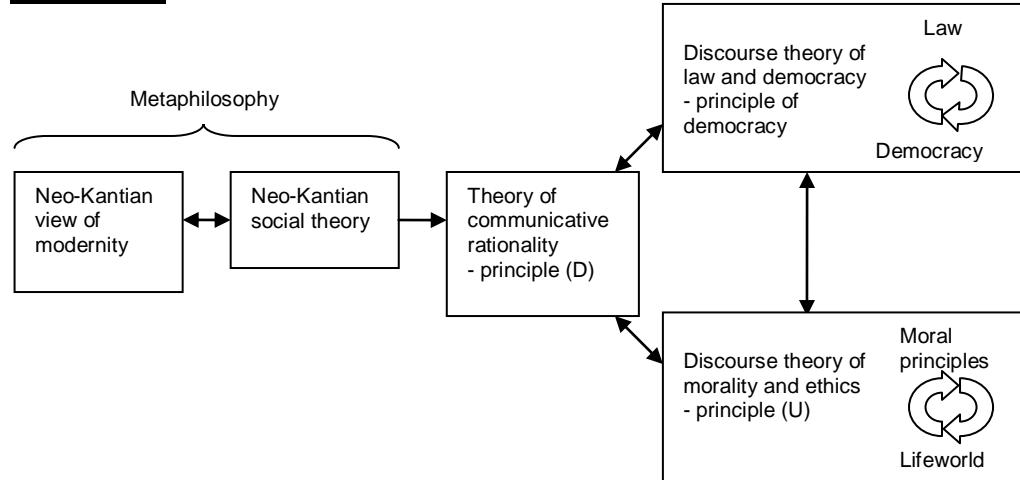


Figure 3. Self-grounding Projects: Rawls’s Hegelianized Kantianism and Habermas’s Kantianized Hegelianism

The problems of modern moral philosophy are intimately connected with Protestantism for Rawls, aiming to ‘establish a basis of moral knowledge independent of church authority and available to the ordinary reasonable and conscientious person’ (LHMP: 8). Rawls thus argues that political liberalism is sharply different from, and

rejects, Enlightenment Liberalism, which historically attacked orthodox Christianity and claims to be suitable for a modern post-Christian age (PL: xl; CP: 611). However, endorsing Waldron's account of the Enlightenment foundations of liberalism, Rawls says political liberalism seeks to satisfy the 'traditional liberal demand to justify the social world in a manner acceptable "at the tribunal of each person's understanding"' (PL: 391 n.28; quoting Waldron 1993: 61). Rawls can be placed within an Enlightenment tradition when this is construed in a broad sense as concerned with justifying the terms of political association in modern democratic societies on terms acceptable to citizens committed to a plurality of worldviews.⁵⁷

3.4 Habermas: Kantian Republicanism as Kantianized Hegelianism

Habermas's project, in contrast to Rawls's, consists of a series of interconnected research programmes. None of these programmes is foundational, their justification depending upon overall coherence between them. Habermas's understanding of modernity is not a separable programme but a 'collection of ideas and assumptions that are woven into all the various programmes' (Finlayson 2005: 63). Important for my purposes in what follows is the way Habermas develops his theory of communicative rationality in the domains of moral theory and legal theory, and I do not discuss in detail the pragmatic theory of meaning that underlies Habermas's theory of communicative rationality.

There is a convergence between Rawls and Habermas in the sense that while Rawls disclaims any connection with a secular Enlightenment tradition, but can be understood as taking forward the Enlightenment 'project', suitably understood, Habermas associates himself with the Enlightenment tradition but does not understand this tradition in Rawls's terms as seeking to transcend, by replacing, religious modes of thought. To understand the Enlightenment as a triumphant secular project would be to reduce its defence to the level of Blumenburg's project of self-assertion. In seeing the Enlightenment as an incomplete self-grounding project, Habermas offers a genealogy in which the critique of the partiality of supposedly universal reason always accompanies its assertion (PDM). The philosophical discourse of modernity is therefore potentially self-correcting, the only solution to the problems of Enlightenment being more Enlightenment.

⁵⁷ What Frazer (2007) argues is a weakness, Rawls's attempt to reconcile the rationalist and sentimental Enlightenment, I regard as a strength.

Habermas understands the movement of thought leading to a detranscendentalized philosophy as a move ‘from Kant to Hegel and back again’ (TaJ: 175-213). This motif can also guide an understanding of Habermas’s interlinked research programmes. First, in his social theory, inspired by Weber, Habermas’s social ontology can be seen as involving a Weberian neo-Kantianism (IO: 70; TCA 1: 339). Kant’s distinction between three forms of reason corresponding to the three critiques – scientific, practical and aesthetic – takes on social form. Hegel historicizes and socializes Kant’s conception of practical reason, but with Weber we return to the Kantian distinction in a socialized form, as traditional sources of legitimation are undermined by social rationalization. The lifeworld becomes divided into resources related to culture, society and personality, a rationalized *sittlichkeit* (TCA 2). Corresponding to this social theory is a modest conception of the role of philosophy. Gone are the claims of ‘first philosophy’ and the assertion of the superiority of theory over practice, with philosophy instead playing the role of stand-in for reconstructive empirical sciences and interpreter between them (MCCA: 1-21).

There is a profound ambivalence in Habermas’ view: ‘Rationalization of the lifeworld means differentiation and condensation at once – a thickening of the floating web of intersubjective threads that simultaneously holds together the even more sharply differentiated components of culture, society, and person’ (PDM: 346). This provides the background for understanding a second move ‘from Kant to Hegel and back again’, this time within Habermas’s whole programme. In the case of both his discourse theory of morality⁵⁸ and his discourse theory of law, normative principles are sustained in an ongoing reflexive project of self-legislation: transcendent Kantian principles defended on a Hegelian immanent methodological basis. Habermas’s comment in relation to the programme of discourse ethics could serve as a motto for his project as a whole. Discourse ethics picks up the Hegelian aspiration to oppose both the abstract universality of justice and the concrete particularism of the common good, but it ‘picks up this basic Hegelian aspiration to redeem it with Kantian means’ (MCCA: 201).

Conclusion

Part I has covered a lot of ground, but its aim has been to prepare the ground upon which more detailed analysis can build. I have examined the RKH line in order to illustrate the joint concern of Rawls and Habermas with the value of autonomy and the self-grounding

⁵⁸ At this stage I use the terms discourse ethics and discourse theory of morality interchangeably, but I explain how Habermas comes to prefer the latter in Chapter 9.3.

problem of postfoundationalist philosophical justification to which this gives rise. Examining Cohen's analytical Marxism, and how this sets the terms of Cohen's move to normative philosophy, has shown Cohen's reasons for rejecting the problem of self-grounding.

On this basis, the positions of Cohen, Rawls and Habermas can be assessed in terms of what I will call their contrasting economies of philosophical ambition, the way they understand the relationship between normative philosophy and social theory. The dual meaning suggested by the expression economy is apt. First, modesty in one of these domains is inversely correlated to ambition in the other, the move from Cohen to Rawls to Habermas involving an increasing translation of traditional philosophical problems into the terms of social theory. In Part II, I will focus on the background social theories of Cohen and Rawls in defending the deep structure of Rawls's constructivism against Cohen's criticisms. Second, in the Rawls-Habermas debate both Rawls and Habermas claim an economy of philosophical ambition, each claiming to propose the more modest approach. Rawls argues that his approach is more modest in being political rather than metaphysical, while Habermas in turn argues that his position is more modest than Rawls's in being procedural rather than substantive, disclaiming the ambition of outlining a blueprint of a just society. But in each case we will need to keep in mind the ambitions that motivate this modesty, since as a result of this philosophical modesty each claim that their theory encompasses that of the other. Rawls argues that Habermas's radically democratic view would be acceptable within public reason, while Habermas argues that Rawls's natural law normativism leaves Rawls's view detached from the facts studied by the sociology of law that Habermas's own theory seeks to encompass. Habermas asks with respect to Dworkin and Luhmann 'Who can embrace whom – the sociologist the legal philosopher, or vice versa? (EFP: 39), and we can ask a similar question with respect to Habermas and Rawls.

With the move from Cohen to Habermas, the argument of this thesis comes full circle. Both Cohen and Habermas reject Rawls's Hegelian strategy of reconciling facts and norms. But their contrasting approaches represent the two leading avenues open to post-Rawlsian political philosophy: condemning the Hegelian remnants in Marxism and resorting to utopian philosophy or seeking in a Hegelian fashion to reconstruct the operation of principles of radical democracy within existing social practices.

PART II. CONSTRUCTIVISM BEYOND REALISM AND UTOPIANISM

4. The Basic Structure and the Division of Justificatory Labour

‘It is . . . a commonplace that no normative political philosophy can get off the ground without making some assumptions about what humans are like’ (John Dupré, quoted in RJE: 232 n.4). Cohen argues that while this may be a commonplace it is mistaken. At least, it would only be true if ‘once political philosophy does get off the ground, it leaves the ground of fact behind’ (RJE: 232 n.4). Likewise, Cohen believes it mistaken to say that ‘it is true of any normative theory . . . that it is linked with certain explanatory theory or theories’ (Charles Taylor, quoted in RJE: 232 n.4). For Cohen, normative theory has no necessary connection with existing social practices, or more generally the facts studied by social theory. In Part II, I will argue against this view by defending Rawls’s constructivism against Cohen’s critique of its fact-dependence. My aim is to reconstruct and defend the deep structure of Rawls’s methodology; engagement with the details and development of Rawls’s position is postponed until Part III. This is important since the common self-grounding project in which Rawls and Habermas are engaged will not get off the ground if Cohen’s critique is successful (or rather their methodological aspiration for normative grounds to remain in touch with the ground of social practices will stand condemned as philosophically illegitimate).

My concern in this chapter is with Cohen’s critique of Rawls’s focus on the basic structure of society as the first subject of justice, a critique that forms the background to Cohen’s later argument against constructivism. Cohen argues that his challenge to constructivism would apply even if constructivists dropped their focus on the basic structure (RJE: 276), but I will argue that for Rawls constructivism and the focus on the basic structure stand and fall together. Indeed, exploring Rawls’s understanding of the basic structure brings into focus the social theory against the background of which Rawls’s constructivist procedure of the original position is employed: ‘choice of the principles of justice presuppose[s] a certain theory of social institutions’ (TJ: 160/138). Focusing on the role of the basic structure, which is unaffected by Rawls’s political turn,⁵⁹ is a fruitful starting point for re-evaluating Rawls’s project. As Pettit points out, ‘an implicit ontology of the people and the relation between the people and the state often shapes how we think in normative terms about politics’ (Pettit 2005: 157). The role of the basic structure illustrates Rawls’s social holism, which, as Charles Taylor observes, tended to be overlooked in the liberal-communitarian debate. Following

⁵⁹ Rawls’s 1978 article ‘The Basic Structure as Subject’ (itself a revision of a 1977 article) dates from long before Rawls’s ‘political turn’ but is reprinted unchanged as Lecture VII of *PL*.

Taylor (1989), we can distinguish between advocacy issues and ontology issues. Rawls is a liberal holist, advocacy of individual rights not precluding an explanatory ontological holism.⁶⁰

In reconstructing Rawls's approach, I begin by considering how Rawls's way of thinking about justice differs to Cohen's. After then setting out Cohen's critique of Rawls's focus on the basic structure of society, I begin my defence of Rawls by assessing Samuel Scheffler's response to Cohen. This is just one aspect of Scheffler's valuable attempt to rebut two philosophical tendencies that have obscured the distinctiveness of Rawls's egalitarianism. The first is an 'egalcentric' narrative of the recent history of political philosophy, a narrative of progress from utilitarian conceptions of equality to liberal ones, and, within the latter, from Rawls's theory of justice to Dworkin's theory of equality of resources (Scheffler 2003a: 16 n.27).⁶¹ But to see Rawls as simply criticising utilitarianism as a mistaken theory of distributive justice overlooks the fact that Rawls's primary objection to utilitarianism is to its mode of justification that fails to respect persons' moral autonomy. (Scheffler 2003b). It is anachronistic to read back into Rawls's conception of social justice a concern with answering the question later posed by Sen (1980), 'Equality of What?'⁶² Closely connected with this egalcentric narrative is the tendency of interpreting Rawls according to an administrative conception of justice, evident in Dworkin's view of equality as a 'virtue of sovereigns' (Scheffler 2003a).⁶³ It is misleading for Dworkin to suggest there is a consensus that political philosophy is concerned with how governments should *treat* their citizens with equal concern and respect, since if Rawls is anything to go by there is no consensus that it should not be principally concerned with the prior question of how citizens should conceive of the terms of their political association. Cohen accepts the egalcentric narrative of political philosophy but rejects the idea of equality as a virtue of sovereigns. Understanding how Rawls's rejection of the latter differs to Cohen's requires recognising his reasons for rejecting the former. Scheffler's analysis points in the right direction, but I will argue that Rawls's 'division of moral labour' between institutions and individuals is ultimately part of a methodological division of justificatory labour in Rawls's constructivism, 'a division of labor between general facts and moral conditions in arriving at conceptions of justice' (TJ: 160/138).

⁶⁰ See also, Scheffler (2003b); Shane O'Neill (1997: 22-24); Pettit (2005, 2006); Audard (2007).

⁶¹ See Kymlicka (1990). On the inadequacies of utilitarianism as an egalitarian view, see Freeman (1994); Scheffler (2003b).

⁶² On why Rawls is not a proto luck egalitarian, see Scheffler (2003a); Daniels (2003); Freeman (2007b: 111-43).

⁶³ See also, Dworkin (2003); Scheffler (2003c).

4.1 Platonic and Aristotelian Ways to Think About Justice

While Cohen acknowledges the depths of his methodological differences with Rawls, he neglects the full ramifications this has for their respective ways of thinking about justice. David Miller traces back the debate between universalist and contextualist ways of thinking about justice to Plato and Aristotle (Miller 1999: 52-53; see also Miller 2002). Since this leaves to one side Aristotle's metaphysics, this is admittedly an oversimplification, but it will prove useful as a way of constructing ideal types for approaching the respective positions of Rawls and Cohen. Cohen accepts a Platonic view that begins from everyday opinions about justice but seeks to correct such opinions through philosophical knowledge (RJE: 291). Justice is ascertained through the exercise of theoretical reason and the standard of justification which is aspired to is the logical proof of mathematics or physics. What justice demands would then be a question of deduction from first principles. By contrast, Aristotelian approaches 'seek to correct common opinion using only methods of argument that common opinion itself endorses' (Miller 1999: 53). Justice is understood through practical reason, which is independent of theoretical reason. While for Aristotle practical reason also proceeds to first principles, these are not fixed and universal, but rather inductive generalisations that are for the most part true. Therefore, while Platonic knowledge transcends everyday opinions – 'justice transcends the facts of the world' (RJE: 291) – Aristotelian knowledge incorporates these opinions.

Cohen connects Marxism with a Platonic tradition in which philosophy 'seeks to realize itself in the world, the proletariat tasked with realizing in the deficient real world the perfect world elaborated by philosophy' (IYE: 97-9). Pursuing this project within the terms of normative philosophy now involves for Cohen unearthing 'the fact-free principle that governs our fact-loaded particular judgements about justice', as until we do so 'we don't know why we think what we think just is just'. Life in Plato's cave can be reordered on normative lines only by being illuminated by the sun of justice: 'facts cast normative light only by reflecting the light that first principles shine on them' (RJE: 267). In a division of labour, socialist philosophers focus on advocacy – what values 'we' ('those of us who remain socialists' (IYE: 77)) are justified in seeking – while socialist economists devise designs for how these values can be implemented institutionally (IYE: 115), the facts of economics and sociology determining 'which implementation packages are feasible' (RJE: 272).

Rawls, on the other hand, presents his understanding of justice as consistent with the traditional Aristotelian view of justice as refraining from *pleonexia*: respecting

persons entitlements by neither taking what belongs to them nor denying them what is due to them (TJ: 10/9-10). While Rawls's view of justice as the first virtue of social institutions, rather than of individual actions, may seem to contradict this self-understanding (Bedau 1978), there is a deeper justification for Rawls's claim. This lies in Aristotle's belief that the possession of a *sense* of what is just and unjust is unique to man and that sharing a common understanding of justice makes a polis. Rawls says by analogy that 'a common understanding of justice as fairness makes a constitutional democracy' (TJ: 243/214). A modern constitutional democracy is not an Aristotelian polis and cannot be a community, but in Hegelian fashion its shared sense of justice, depending upon the formative influence of social institutions, still allows for the realization of civic friendship (TJ: 5/5). I focus on the importance of Rawls's stability argument in Part III, but we may note here that if Hegel's moral psychology seeks to reconcile Kant and Aristotle (DeLue 1980), Rawls's project is in this respect also Hegelian. Rawls describes our sense of justice developing according to an Aristotelian Principle, but his Kantian congruence argument seeks to show that the reflective endorsement of Kantian principles of right is congruent with, indeed the culmination of, individuals' conceptions of the good. Unlike in Aristotle there is no 'dominant end' within a rational plan of life, but rather an inclusive end in which principles of the right structure, but do not transcend, conceptions of the good (TJ: 528/463).

While Rawls's understanding of justice is the focus of this chapter, it dovetails with Rawls's understanding of justification addressed in the next chapter. Rawls begins from the assumptions of Miller's 'Aristotelian', or real feasibility, approach, focused on common sense shared understandings, but he does not accept that this is where philosophical reflection must end up (PL: 44 n.47). If Rawls's conception of justice begins from an 'Aristotelian' basis but has a Hegelian twist, so does his conception of justification. Reflective equilibrium 'goes back in its essentials to Aristotle's procedure in the *Nicomachean Ethics*' (TJ: 51 n.26/45 n.26).⁶⁴ Justification presumes a clash of views between persons or within ourselves, drawing attention to the Socratic aspect of moral theory, its concern with self-knowledge. But the corollary of the Socratic nature of moral theory is that 'we may want to change our present considered judgements once their regulative principles are brought to light' (TJ: 49/omitted). Justification proceeds from a consensus – that is why it differs from 'mere proof' (TJ: 581/508) – but attempts to extend and deepen that consensus. Miller's invocation of Aristotelian theoretical modesty is misleading inasmuch as Aristotle thought that on the basis of induction from

⁶⁴ See Kelly (2000: 239-40); Hooker (2000: 9 n.9). Following Schneewind (1963), Rawls takes Sidgwick to be broadly committed to systematizing common sense rather than replacing it with intuitionistic principles (TJ: 51/45). See also Schneewind (1974), but cf. Singer (1974).

accepted beliefs one could arrive at true first principles. Rawls's employment of the original position maintains an aspiration for arriving at structuring principles, but these first principles are contingent not necessary truths (TJ: 578/506), Rawls disagreeing with Aristotle in arguing for a circular rather than foundational form of justification (Føllesdal 2005: 170-71). Justification by reflective equilibrium is a matter of systematicity, of 'everything fitting together into one coherent view' (TJ: 579/507). In its Socratic nature, Rawls's 'dialectic' of reflective equilibrium can be compared to Hegel's dialectic to the following limited extent. Both can be seen as providing a 'metalanguage' in which problems posed in terms of an object-language are resolved, but 'preserved' rather than simply being transcended, principles of justice accounting for our considered judgements in Rawls's conception (Schwarzenbach 1991: 546). The original position is only ultimately justified if it corrects common opinion using only methods of argument that common opinion itself endorses, for Rawls claims that the conditions on the choice of principles represented by the veil of ignorance 'are ones that we do in fact accept', or 'can be persuaded to do so by philosophical considerations' (TJ: 587/514).

Cohen argues that *TJ* promised to tell us 'what justice is' (RJE: 305), but there are already grounds for believing that this can only result from reading *TJ* through the lens of a methodological paradigm that Rawls rejects; perhaps, as Coleridge suggested, one is born either a Platonist or an Aristotelian. Rawls famously begins *TJ* by arguing that 'justice is the first virtue of social institutions, as truth is of systems of thought' (TJ: 3/3). Cohen argues, by contrast, that the relationship between social institutions and justice is closer to the relationship between utterances and truth than it is to the relationship between systems of thought and truth. Just as 'the truth of a statement is neither a necessary nor a sufficient condition of its justifiable utterance', so some justice cannot be implemented institutionally (RJE: 303-4). But this simply begs the question, depending upon Cohen's governing distinction between justice and rules of regulation. In 'taking men as they are and laws as they might be', Rawls's focus on the basic structure is intrinsically related to his holistic Quinean understanding of justification through reflective equilibrium (J. Cohen 2002). In his equation of justice and the first virtue of social institutions, Rawls is not claiming synonymy. As Quine put it, 'If we paraphrase a sentence to resolve ambiguity, what we seek is not a synonymous sentence, but one that is more informative by dint of resisting some alternative interpretations' (Quine 2001 [1960]: 159). Cohen attempts to rescue the concept of justice from the 'neologizing' sense in which it is applied by constructivists (RJE: 3, 303), but for Rawls philosophy is not the ahistorical pursuit of 'an elusive virtue discussed for a few thousand years by philosophers' (RJE: 304). It must instead be concerned with, and may

legitimately aspire to, conceptual revision. Justice will be the first virtue of social institutions if Rawls is correct that regarding justice in this way systematically reorders our considered judgements about justice in accordance with principles we accept. As Rawls put it in arguing against traditional analytic approaches to ethics, ‘morals is not like physics: it is not a matter of ingenious discovery but of noticing lots of obvious things and keeping them all in reasonable balance at the same time’ (Rawls 1951: 579-80; see also TJ: 587/514).

4.2 Justice of the Basic Structure and Justice of Personal Behaviour

Cohen’s ‘Platonic’ argument for correspondence between justice in the ordering of society and justice in personal behaviour begins from a critique of Rawls’s acceptance of material inequalities resulting from economic incentives. Cohen takes Rawls’s difference principle to be an attempt to justify income differentials when they are necessary to provide incentives for more productive workers to work to their productive potential in a way that benefits the worst off in society. But in what sense, Cohen asks, are such incentive payments ‘necessary’? If they are strictly necessary, ‘necessary, that is, apart from people’s chosen intentions’ (RJE: 69), then they would be justified as a matter of principle. This would support a strict reading of the difference principle. However, if they are only necessary ‘given that talented producers operate as self-interested market maximizers’, then they will only be justified as a pragmatic compromise with existing facts, ‘a principle for handling people’s injustice’ (RJE: 69, 84). This would support a lax reading of the difference principle, appropriate as a principle of public policy but not as a principle of justice.

Cohen argues that Rawls’s remarks about the difference principle applying to a well-ordered society in which everyone knows, affirms, and acts from principles of justice commits him to the strict reading. But there remains an ‘unresolved tension in the Rawlsian architectonic’ between two conceptions of social relationships: a bargaining conception in which the talented will only work to their productive potential in return for incentive payments, and a community conception in which incentive payments are unnecessary because the talented have fully internalized the demands of the difference principle (RJE: 82; see also Barry 1989). Accordingly, Cohen argues that a fully just society will be characterised by an egalitarian ethos. To be sure, individuals have what Scheffler calls an ‘agent-centred prerogative’, ‘*a right to pursue self-interest to some reasonable extent*’ (RJE: 61), but justice in the full sense requires a culture of justice that

inculcates respect for the strict difference principle across a full range of market and other personal behaviours.

My primary concern is not with the details of this argument, a subject that has generated a significant literature, although I will return to the incentives argument later.⁶⁵ My interest is rather in the methodological significance of Cohen's critique insofar as it relates to Rawls's focus on the basic structure. The basic structure enters the picture in the form of a Rawlsian counter-argument against the requirement of an egalitarian ethos, the 'basic structure objection', that Cohen seeks to refute. According to this objection, principles of justice apply only to the coercive basic structure of society and not to the choices that individuals make within this coercive structure. But Cohen argues that this restriction in the scope of justice is inconsistent with Rawls's view that a well-ordered society is one characterised by fraternity, where the worst off can accept their position with dignity and persons fully realize their moral nature by acting from principles of justice. More fundamentally, it is not clear whether 'the Rawlsian basic structure includes only coercive aspects of the social order or also conventions and usages that are deeply entrenched but not legally or literally coercive' (RJE: 125). If the former is the case then significant areas of social life, notably the family, will be beyond the purview of principles of justice. Logical consistency demands that Rawlsians recognise that the reasons they have for applying principles of justice to coercive institutions – that the effects of such institutions are 'profound and present from the start' (TJ: 7/7) – commit them to applying these principles to personal behaviour. The family, for example, meets both these criteria. However, to include the family within the basic structure would have fatal results for Rawls's position:

[O]nce the line is crossed, from coercive ordering to the noncoercive ordering of society by rules and conventions of accepted practice, then the ambit of justice can no longer exclude chosen behavior, since, at least in certain cases, the prescriptions that constitute informal structure [as with the family] are bound up with the choices that people customarily make (RJE: 134).

This means that for Cohen 'the personal is political' (RJE: 116). Personal choices beyond the scope of law are properly matters of justice.

Murphy generalises Cohen's point in the form of a methodological distinction between dualism and monism.⁶⁶ The dualist 'discontinuity' view accepts, while the

⁶⁵ See Williams (1998); Wolff (1998); Estlund (1998b); Smith (1998); Murphy (1999); Pogge (2000); Farrelly (2000, 2005); J. Cohen (2002); Tan (2004); Voorhoeve (2005); Baynes (2006).

⁶⁶ It should be noted that Murphy's position differs to Cohen's to the extent that it is derived from a general utilitarian standpoint whereas Cohen aspires to an internal critique of Rawls (RJE: 397). Murphrey's monism is neutral with respect to the pluralism of fundamental principles that Cohen

monist ‘continuity’ view rejects, the proposition that ‘the two practical problems of institutional design and personal conduct require, at the fundamental level, two different kinds of practical principle’ (Murphy 2002: 71; Murphy 1999: 254). Dualism is understood by Murphy as a response to ‘the problem of the unreasonable demands [egalitarian] theories of justice may potentially impose on some people’ (Murphy 1999: 255). This, it is argued, motivates Rawls’s idea of an

institutional division of labour between the basic structure and the rules applying directly to individuals and associations and to be followed by them in particular transactions. If this division of labour can be established, individuals and associations are then *left free to advance their ends more effectively within the framework of the basic structure*, secure in the knowledge that elsewhere in the social system the necessary corrections to preserve background justice are being made (PL: 268-69, emphasis added).⁶⁷

Cohen and Murphy accept the conceptual framework of, and take Rawls to be responding to, Nagel’s moral division of labour (Murphy 2002: 70-73). According to the purported ‘Rawls-Nagel division of labour’ (Julius 2003: 326-27), institutions ‘take the business of securing justice off people’s plates in their day-to-day lives’. Thus we can ‘eat our cake and have it too’, achieving ‘egalitarian aims without making ourselves miserable in the process’ (Murphy 1999: 258).⁶⁸ While Cohen recognises that Rawls differs from Nagel in seeing individuals as having a natural duty to uphold or foster just institutions, Cohen assumes the terms of Nagel’s framework, what, following Sidgwick, we could call a dualism of practical reason between egoism and benevolence (RJE: 9).

Scheffler’s response to Cohen is valuable in emphasising that while the idea of a division of moral labour⁶⁹ is central to Nagel’s work, Rawls has other motivations. Nagel’s idea of a division of moral labour is based on a division between the personal and the impersonal standpoints within each individual. We can ‘externalize through social institutions the most impartial requirements of the impersonal standpoint’ in such a way that ‘our support of those institutions depends on the fact that they answer to the demands of a very important part of ourselves’ (Nagel 1991: 53). This sets the terms for Cohen’s argument, terms that not surprisingly are in their turn endorsed by Nagel (2000). But in seeking to represent the point of view of reasons we can share, Rawls’s

affirms, as monism says only that *if* one decides to apply principles of institutional design *then* one should apply the same principles to individuals (RJE: 394 n.46).

⁶⁷ In another formulation the italicised phrase is rendered as ‘left free to advance their (permissible) ends within the framework of the basic structure’ (JF: 54).

⁶⁸ If true this would be a marked change from Rawls’s undergraduate view that ‘to use the method of legalism [to bring about community] is to thrust us further into aloneness and separation’ (MSF: 227).

⁶⁹ I follow Scheffler in preferring this formulation.

constructivism rejects Nagel's idea of 'the impersonal point of view' (PL: 116). There is for Rawls 'no discontinuity at the level of moral psychology that corresponds to the distinction between institutional and non-institutional values' (Scheffler 2005: 247).

What, then, are Rawls's motivations? Scheffler offers an alternative account of the connection between the institutional division of labour and the division of moral labour. The rationale of the latter is to 'allow for the plurality of values and principles' and 'clarify the limited scope of Rawls's project' (Scheffler 2006: 106). For Scheffler, egalitarian liberalism's foundational moral commitment is to respecting moral pluralism, exhibiting philosophical modesty in seeking to reconcile 'the perceived tension between the values of personal life and the values of political morality' by giving each of these dimensions of our moral life their due (Scheffler 2005: 235).⁷⁰ The institutional division of labour is

not about pluralism or methodological modesty, but rather that, given the feasibility constraints that apply to the rules regulating individual conduct, the institutional forms required to ensure background justice will have to go beyond and help fix the content of those rules (Scheffler 2006: 106-7).

Scheffler's division of moral labour is fundamentally distinct from the institutional division of labour: preserving the background justice of economic transactions 'has nothing to do with pluralism' (Scheffler 2005: 240). However, I will argue that separating economic and social dimensions in this manner is unpersuasive both in itself and as an interpretation of Rawls. On Scheffler's view we first respect moral pluralism and then allocate the task of fulfilling values that go beyond our immediate personal and social lives to economic institutions. But for Rawls the choice of social institutions must 'be made on moral and political as well as on economic grounds' (TJ: 260/229). Rawls's reasons for taking the basic subject as the first subject of justice undermine Scheffler's distinction.

It is true that the first reason Rawls gives for focusing on the basic structure is the economic goal of maintaining economic background justice. (JF: 54). Contrary to Robert Nozick's suggestion that Rawls proposes end-result patterned principles, Rawls's is an 'ideal historical process view'.⁷¹ Principles of justice give a regulative structure to the arrangement of social institutions, just distributions being those that result from the operation of just regulative rules, a matter of pure procedural justice. However, the second reason that Rawls focuses on the basic structure stands in marked contrast to Scheffler's interpretation. This is its educative role, an aspect of the role of social

⁷⁰ See also Scheffler (1994, 2001).

⁷¹ Kaufman (2004) refutes the myth that Rawls proposes patterned principles of justice.

institutions that Rawls finds in Hegel (LHMP: 353, 367). As a system of social and economic institutions, the basic structure is not only an arrangement that satisfies given desires and aspirations, but also arouses further desires and aspirations in the future (JF: 56; see also TJ: 259/229). Contrary to what Scheffler suggests, for Rawls the functioning of free social institutions does not simply respect the fact of pluralism, but constitutes this pluralism in the first place. More than just playing the ‘public role of educating citizens to a conception of themselves as free and equal’, a properly regulated basic structure encourages in citizens ‘attitudes of optimism and confidence in their future, and a sense of being treated fairly’ (JF: 57). Understanding the institutional division of labour and the division of moral labour as part of a holistic view will therefore require understanding their place within an overall division of justificatory labour.

4.3 From the Division of Moral Labour to the Division of Justificatory Labour

Notwithstanding the limitations of Scheffler’s view, his observation that egalitarian liberalism has developed in dialogue with utilitarianism and libertarianism is a helpful starting point (Scheffler 2005: 237). Justice as fairness agrees with libertarianism in condemning the utilitarian application of general moral principles to both institutional organisation and individual behaviour for failing to respect individual moral autonomy. But while for libertarianism ‘micro-level norms define the content and limits of macro-level norms’, justice as fairness shares with utilitarianism a holism about distributive justice. The justice of an individual’s holdings is a function of principles that regulate the justice of a social distribution as a whole (Scheffler 2005: 234, 249). Theorists like Nozick and Cohen who conceive of autonomy in terms of self-ownership confuse Rawls’s opposition to utilitarian holism with opposition to holism per se:

For them, constructiveness, systematicity, and holism may all be symptomatic of a failure to attach sufficient moral importance to the separateness of persons. If so, however, then their ultimate concern is not the same as [Rawls’s], even if it can be expressed in the same words (Scheffler 2003b: 450-51).

The problem is, then, how to reconstruct economic and social institutions in a manner that respects individual autonomy and social pluralism because it promotes political autonomy and affirms reasonable pluralism.

It is unfortunate that, while not inconsistent with his later definition of the basic structure, the definition that Rawls gives in *TJ*, and which has become canonical (‘the way in which the major social institutions distribute fundamental rights and duties and

determine the division of advantages from social cooperation' (TJ: 7/6)) is not the most informative. Rawls's definition of the basic structure in *PL* is more revealing. The basic structure of society is said to consist of 'society's main political, social, and economic institutions, *and* how they fit together into one unified system of social cooperation from one generation to the next' (PL: 11, emphasis added).⁷² Notice the dual aspect of Rawls's definition, the basic structure of society being the 'basic structure of the social system as a whole' (TJ: 57/50). This suggests two different standpoints one may adopt in relation to the basic structure. On the one hand, society is a *social system* that partly determines who we are and who we want to be. But we are not simply determined by society. We can reform society by reforming the basic structure of the social system. As we will see in Chapter 6, Rawls's understanding of the basic structure extends to the political domain his conception of social practices. The basic structure is a social practice of social practices, in the ideal form of a well-ordered society becoming a social union of social unions. One cannot *observe* how political and social institutions 'hang together as one system of cooperation' (PL: 202). Such a system is constituted by the philosophical *ideas* we employ in conceiving of society. Whether or not it inspires Rawls's terminology, the basic structure can be seen as combining base and superstructure in Marx's sense. Unlike Cohen's interpretation of this distinction, in which the base is independent of, and determines, the superstructure, Rawls develops something akin to a Hegelian-Marxist conception in which base and superstructure are internally related, the superstructure determining the base, as well as vice versa.

Nozick asks 'In virtue of what features of the basic structure, features not possessed by microcases, do special moral principles apply that would be unacceptable elsewhere?' (Nozick 1974: 205). On Rawls's view, while major social institutions may have developed historically as a result of a quantitative increase in social interactions, the idea of the basic structure is an emergent property of the different qualitative attitude we take towards these institutions. This is when we no longer regard society as a hierarchical order and see persons as capable of achieving autonomy as a result of ordering social institutions in accordance with the priority of liberty.

While a social practice standpoint on the basic structure is motivated by similar reformist concerns to those of classical utilitarianism, it lacks utilitarianism's general scope.

[I]f we assume that the correct regulative principle for anything depends on the nature of that thing, and that the plurality of distinct persons with separate systems of ends is an essential feature of human societies, we

⁷² See also CP: 97, 156; JF: 10; PL: 258.

should not expect the principles of social choice to be utilitarian (TJ: 29/25; see also PL: 262).

Rawls is critical of utilitarianism as much as he is of any comprehensive doctrine that conceives of a well-ordered society as governed by a single doctrine (PL: 42-43). There is therefore a sense in which libertarianism provides a corrective to the utilitarian view. But ‘taking men as they are’ does not mean accepting libertarianism’s social atomism. Conceiving of society as a voluntary association of autonomous individuals ‘overlooks the prior and fundamental role of basic institutions in establishing a social world within which alone we can develop with care, nurture, and education, and no little good fortune, into free and equal citizens’ (PL: 43). Libertarianism fails to grasp the social preconditions of autonomy. *TJ*, on the other hand, ‘follows Hegel . . . when it takes the basic structure of society as the *first* subject of justice. People start as rooted in society and the first principles of justice they select are to apply to the basic structure’ (LHMP: 366). Justice as fairness can successfully reply to Hegel’s criticisms of social contract theory because it has already taken this Hegelian point on board (PL: 285-88).

The second perspective on the basic structure is, then, one where the basic or background institutions of the basic structure – the political constitution, and associated form of government and legal system, the systems of property and the market, and the family in some form (Freeman 2007a: 101) – form our social world.⁷³ Institutions are something that we are born into (and in which Rawls stipulates we are to be thought of as leading a complete life). If the social practice standpoint is a top-down view from the standpoint of the original position, the social world standpoint is the bottom-up standpoint of socialized individuals. Basic institutions are akin to a hermeneutical horizon against the background of which our conceptions of ourselves as persons develop. All forms of social cooperation, and the explicit and implicit agreements that they involve, exist within ‘some more or less clearly specified situation embedded within the background institutions of the basic structure’ (PL: 23). The original position must ‘extend the idea of agreement [embedded within the background framework of the basic structure] to this background framework itself’ (PL: 23). The veil of ignorance is needed to ‘abstract from and not be affected by the contingencies of the social world’ (PL: 23). Justice demands that the rules of the social ‘game’ be those that could have been agreed to in an initial situation of fairness. Rawls does not therefore assume that all facts about persons and society can be reformed:

⁷³ See PL: 11, 23, 41, 43, 77, 229, 262-71. While Hegel does not use the German equivalent *die soziale Welt*, Hardimon (1994: 16) employs the term social world as equivalent to Hegel’s ‘ethical world’. As such it is ‘very closely related to Rawls’s idea of the “basic structure”’.

The natural distribution is neither just nor unjust; nor is it unjust that persons are born into society at some particular position. These are simply natural facts. What is just and unjust is the way that institutions deal with these facts (TJ: 102/87).⁷⁴

Utilitarianism adopts a technocratic social practice standpoint while libertarianism treats inequality as a natural and unchangeable fact of our social world; justice as fairness seeks to avoid the errors of both.

The distinction between the social practice and social world standpoints on the basic structure becomes more concrete when one considers it in relation to Rawls's social ontology of the public political forum and the background culture of civil society. From the standpoint of civil society, the basic structure is a background that we are *subject to*, our social world, whereas from the standpoint of the public political forum it is reflected upon as an *object of* reform: principles of justice cohere social practices into an overall social practice of cooperation between free and equal persons. In a manner reminiscent of Hegel's social ontology in the *Philosophy of Right*, within the background culture of civil society persons are seen as members of groups and associations, 'churches and universities, learned and scientific societies, and clubs and teams' (PL: 14, 382-85). But beyond just consisting of associations, civil society is a *point of view* from which conceptions of justice are discussed, the point of view of nonpublic reasons. The public political forum, on the other hand, is the point of view of public reason. Here Rawls departs from Hegel in rejecting the idea of representation through estates, seeing citizens as capable of adopting a 'Kantian' standpoint of individual autonomy, although this is political autonomy, not Kantian comprehensive moral autonomy.

None of this is very helpful if we are looking to Rawls to provide a criterion to 'design a blueprint of ideal institutions' (Pogge 1989: 12), since Rawls's characterisation of the basic structure 'does not provide a sharp definition, or criterion, from which we can tell what social arrangements, or aspects thereof, belong to it' (JF: 12). But Rawls is seeking to set out a framework of thought that guides our practical thinking, not a theoretical blueprint. This framework is not without practical implications, though. The free market economy must be regulated to ensure it does not encroach on the autonomy of, and therefore the autonomy of individuals within, the public political forum. Rawls's favoured regime of property-owning democracy must 'put in the hands of citizens generally, and not only of a few, sufficient productive means for them to be fully cooperating members of society on a footing of equality' if it is to 'realize in the basic

⁷⁴ Rawls rejects the 'principle of redress', that *all* undeserved inequalities, natural and social, demand compensation (TJ: 101/86).

institutions the idea of society as a fair system of cooperation between citizens regarded as free and equal (JF: 140).

When it comes to the question of justice in and of the family, Rawls's reference to aspects of the basic structure is instructive.⁷⁵ The family is part of the basic structure insofar as its role is the 'orderly production and reproduction of society and its culture from one generation to the next' (CP: 595). We don't want principles of justice to apply directly to the internal life of the family, since no particular form of the family is required to fulfil this task (JF: 163), but the family is *indirectly* bound by principles of justice since the freedom and equality of women and the equality of children as future citizens are guaranteed by political principles. As Scheffler argues, while Cohen treats the coercive/non-coercive distinction as a categorical distinction between two types of *institution*, it is more plausible to treat it as a distinction between coercive and non-coercive *structures*, both of which may be found within the same institution (Scheffler 2006: 121-22). Criminal and family law, for example, constitutes coercive structure within the institution of the family. We should remember, however, that Rawls speaks of coercion in a Rousseauian sense. What makes aspects of the family part of the basic structure is that their regulation is properly subject to principles publicly justifiable in accordance with a conception of persons as free and equal citizens, or of children as future citizens. What we are ultimately looking for is a 'division of labor between political and social institutions, on the one hand, and civic society with its many and diverse associations', and the family, on the other (LP: 127), one which is motivated by a 'division of labour between different kinds of principles' (CP: 597-98). The institutional division of labour embodies a division of moral labour because of the need for a *justificatory* division of labour between public and nonpublic reason.⁷⁶

4.4 Autonomously Realizing Autonomy

As Rawls summarises the nature of political constructivism:

[A] constructivist view such as justice as fairness, and more general liberal ideas, do not begin from universal first principles having authority in all cases. In justice as fairness the principles of justice for the basic structure of society are not suitable as fully general principles: They do not apply to all subjects, not to churches and universities, or to the basic structures of all

⁷⁵ Rawls clarifies this in response to Okin (1989).

⁷⁶ Laden (2003: 385) uses the phrase 'division of justificatory labor' in reference to the importance of 'the position to and from which certain reasons are offered'. My analysis gives this a broader application. On Rawls's idea of a division of labour in justification, see also Macedo (1990); Weithman (1994); HR 2: 85, 94. See more generally Bohman (1999).

societies, or to the law of peoples. Rather, they are constructed by way of a reasonable procedure in which rational parties adopt principles of justice for each kind of subject as it arises. Typically a constructivist doctrine proceeds by taking up a series of subjects, starting, say, with principles of political justice for the basic structure of a closed and self-contained democratic society. That done, it then works forward to principles for the claims of future generations, outward to principles for the law of peoples, and inward to principles for special social questions . . . Its authority rests on the principles and conceptions of practical reason, but always on these as suitably adjusted to apply to different subjects as they arise in sequence; and always assuming as well that these principles are endorsed on due reflection by the reasonable agents to whom the corresponding principles apply (CP: 532-33).

Political constructivism seeks to realize ‘a conception of practical reason in the basic institutions of society’ (Baynes 2006: 193; see LHPP: 237). This involves a justificatory division of labour between a Kantian procedure of practical reason representing the autonomy of free and equal persons within a Hegelian holistic social ontology of the background culture of civil society that reproduces society as an ongoing system of social cooperation. The

principles and ideas, and the laws they inform, that structure and regulate th[e] relations [between citizens] are not imposed by natural or circumstantial forces beyond citizens’ control, but can be represented as a product of the free use of their reason when exercised in ways that best express their nature as free and equal and reasonable and rational persons (Freeman 2007a: 461).

While sharing a utilitarian commitment to reform of institutions in accordance with a moral conception, Rawls argues that utilitarianism fails to respect individuals’ capacity for political autonomy. ‘Autonomy is not a value advocated in contrast to . . . utility. Autonomy is a practice, a mode of being, of acting and thinking, which should reverberate in the theoretical conception itself’ (Audard 2007: 53). Justice must therefore be capable of being justified justly, so that citizens may realize political autonomy autonomously. Hence the importance of principles of justice not applying directly to the internal life of nonpublic institutions. Society does not consist of separate ontological domains, public reason contrasting not with private reason, but with the social reason of the associations of the background culture and the domestic reason of families (PL: 220 n.7). Public and nonpublic spheres ‘fall out from the content and application of the conception of justice and its principles’, such that if the ‘so-called private space is alleged to be a space exempt from justice, then there is no such thing’ (CP: 599).

Although for Rawls the right is prior to the good, the two are complementary:

just institutions and the political virtues would serve no purpose – would have no point – unless those institutions not only permitted but also sustained conceptions of the good (associated with comprehensive doctrines) that citizens can affirm as worthy of their full allegiance. A conception of political justice must contain within itself sufficient space, as it were, for many ways of life that can gain devoted support. If it cannot do this, that conception will lack support and be unstable. In a phrase, the just draws the limit, the good shows the point (JF: 140-41).

In a circular process, public principles of justice are endorsed on the basis of conceptions of the good formed within nonpublic spheres demarcated by public principles. According to Anderson, the way that Rawls ‘simultaneously appeals to the natural outlook of a democratic society to found his conception of the person, and to his conception of the person to found the basic structure of a democratic society’ means that ‘[i]n a vicious circle, public arrangements are deduced from personal capacities defined as adapted to public arrangements’ (Anderson 2005: 108). But institutions allow for a motivational bootstrapping (Dietsch 2005), a virtuous spiral of social learning.⁷⁷ Justice as fairness addresses itself to reasonable persons, who ‘desire for its own sake a social world’ – a background of just institutions – ‘in which they, as free and equal, can cooperate with others on terms all can accept’ (PL: 50; Freeman 2007b: 34). This may be compared to Elster’s idea of rational precommitment, particularly as this applies to the justification of constitutionalism (Freeman 2007b: 33). In desiring to live according to principles that all can share, Rawlsian reasonable persons jointly precommit themselves to public principles that they desire as regulative of their personal pursuit of their conceptions of the good, principles that are justified by their reflectively cohering with these conceptions of the good. The desire to do right is not simply the desire to do what is right for its own sake, but a contractualist desire that ‘presumes an account of the subject matter of morality as part of its content’ (Freeman 1991: 292), in this case a conception-dependent desire to realize one’s status as a free and equal person in a just social world.

According to Rawls’s moral psychology of the reasonable, with its idea of reciprocity, ‘the reasonable generates itself and answers itself in kind’, since ‘the idea of reciprocity appears both as a principle giving its content and as a disposition to answer in kind’ (JF: 196). It is as citizens educated to understand ourselves as such by the political institutions of our social world that we develop a capacity for reasonableness. Reasonable institutions, which form our public social world, generate reasonable citizens, whose capacity for reasonableness sustains these institutions and is employed

⁷⁷ This possibility is recognised but criticised by Haslett (1987). Cf. Aronovitch (1996).

by them in continuing to reshape and reform them in the direction of an ideal of well-orderedness. ‘We are, in cooperatively devising and maintaining basic social institutions, indirectly shaping ourselves and defining practical reason’ (Freeman 2007b: 43). The resemblance of this to Hegel’s idea that norms are self-legislated insofar as ‘spirit is a result of itself’ becomes more defensible when we view Hegel, and Rawls, from the perspective of a post-Kantian Hegel. On such a view, spirit is ‘a kind of norm’, ‘an achieved form of individual and collective mindedness, and institutionally embodied recognitive relations’ (Pippin 2008: 60, 62, 39). In social terms, principles of justice both create and sustain the conditions needed to realize the status of persons as free and equal and their fundamental interest in cooperation based on mutual respect, or, we could say, recognition (Freeman 2007b: 34).

Constructivism is a ‘view about how the structure and content of the soundest moral doctrine *will* look once it is laid out after due critical reflection’ where ‘by full reflection is not meant perfect reflection at the end of time, but such increasing critical reflection as might be achieved by a tradition of thought from one generation to the next’ (LHMP: 274). It is through a conception of the educative role of social institutions that Rawls seeks to maintain a link with the Platonic aspirations of political philosophy, but in a form radically reconceived for a democratic social world. As Yeats put it, education is not filling a bucket, but lighting a fire.

4.5 Justificatory Community and Individual Self-realization

While we have seen that Rawls rejects Nagel’s pessimism about reconciling the dualism of practical reason, more needs to be said about how public and nonpublic standpoints are ideally reconciled in Rawls’s idea of a well-ordered society. Also, while Cohen may be wrong to see the role of the basic structure as unburdening individuals of moral demands, we are yet to see how Rawls’s division of justificatory labour helps to explain the role of the difference principle and Rawls’s reason for accepting economic incentives. I will argue that both questions can be answered in relation to a Rousseauian conception of a well-ordered society and the counterfactual imagination that is its methodological corollary.

Rawls’s conception of a well-ordered society may be instructively contrasted with Cohen’s conception of justificatory community. Cohen, it will be recalled, sees socialism as combining a principle of luck egalitarian justice and a principle of community. While working within a Marxian framework, Cohen conceived of community as instrumental to the achievement of persons’ independently specified

goals, a ‘concert of mutually supporting self-fulfilments, in which no one takes promoting the fulfilment of others as any kind of obligation’ (SFE: 123):

[I]magine a jazz band each player in which seeks his own fulfilment as a musician. Though basically interested in his own fulfilment, and not in that of the band as a whole, or of his fellow musicians taken severally, he nevertheless fulfils himself only to the extent that each of the others also does so, and the same holds for each of them (SFE: 122).⁷⁸

While socialists could once hope that material abundance would make wage incentives unnecessary (SFE: 126), for Cohen the fact that such abundance no longer seems feasible does not undermine a Marxian conception of community. Since Marx’s optimism about material possibility was the obverse of his pessimism about social possibility, we simply need to reverse this relation (SFE: 132). Cohen remains committed to abolishing the dualism of state and society, but on his utopian socialist conception the withering away of the state arises from a revolution in personal behaviour (RJE: 1-2). For Cohen, social change under historical materialism depended upon material facts that ‘are asocial in that no information about social *structure* enters into their formulation’, where facts about social structure refer to facts about the rights and effective powers persons hold in relation to other persons (HLF: 83). Social change under utopian socialism, on the other hand, depends upon norms that are asocial in that no information about the facts of persons’ relationships to each other within the basic structure enters into their formulation. Cohen sees Rawls’s principles of justice along Nozickian lines as end-result patterned principles, but if he ultimately rejects a Nozickian principle of self-ownership he remains committed as a philosophical ideal to a conception of autonomy as absence of constraint or coercion by social structures.

The alternative to Marx’s optimism about material abundance is, then, optimism about persons voluntarily committing themselves to principles of justice. Justificatory community is achieved when the presentation of arguments or policies passes an interpersonal test. This ‘asks whether the argument could serve as a justification of a mooted policy when uttered by any member of society to any other member’ (RJE: 42). Policies that meet the test have a comprehensive justification. Such a society would be a concert of mutually supporting self-fulfilments, in which, subject to an agent-centred prerogative, everyone takes promoting the fulfilment of others not as an obligation, but as fulfilling the demands of justice.

⁷⁸ For criticism, see Graham (1990). In the hands of Eagleton (2007: 171-73), Cohen’s image of a jazz band comes as close as we can get to ‘the meaning of life’.

Is this really optimism about *social* possibility? Cohen rejects a conception of the social individual as an alternative to material abundance, believing it better to ‘reject the bourgeois normative principle of self-ownership’ than to accept it and rely on an ‘extravagant’ degree of socialization (SFE: 134-45). But extravagant or not, this is Rawls’s preferred option. Rawls’s developed understanding of the problem to be resolved in the formation of a well-ordered society does not depart in essence from his undergraduate view that ‘the chief problem of politics is to work out some scheme of social arrangements which can so harness human sin as to make the natural correlates of community and personality possible’ (MSF: 127-28). In its later Rousseauian form, Rousseau’s remark that ‘[t]he private will tends by its nature towards preferences, and the general will toward equality’ becomes ‘an ancestor of the first reason why, in justice as fairness, the basic structure is taken as the primary subject of justice’ (LHPP: 234). Rawls employs a musical analogy to describe how a Rousseauian idea of a well-ordered society achieves social unity in diversity.⁷⁹ He compares the idea of a well-ordered society as a social union of social unions to a

group of musicians every one of whom could have trained himself to play equally well as the others any instrument in the orchestra, but who each have by a kind of tacit agreement set out to perfect their skills on the one they have chosen so as to realize the powers of all in their joint performances (TJ: 523 n.4/459 n.4; see also JF: 76, PL: 320-24).

By contrast, in Marx’s vision of full communism, ‘if we were musicians we might want to take turns playing all the instruments in the orchestra’ (LHPP: 369).

Much could be said to defend Rawls’s egalitarianism and its conception of a property-owning democracy against Cohen’s criticisms. But when all is said and done, the question remains as to why Rawls should countenance material incentives. In reply, one can note first that while justice is the *first* virtue of social institutions it is not the only virtue. A conception of justice for Rawls should incorporate the virtue of efficiency, the most convincing efficiency-based argument for social inequalities being more that competition is an enabling mechanism that rewards innovation rather than its

⁷⁹ In *TJ*, Rawls attributes this view to Wilhelm von Humboldt (1993) and Kant (TJ: 523 n.4/459 n.4). He does not mention Rousseau or Hegel. However, Humboldt’s political theory may be described as Rousseauian (Burrow 1993: 1), and when Rawls describes how for Rousseau citizens achieve moral freedom through recognition of others he remarks: ‘Think of how the trained powers of musicians reach their fullest fruition only when exercised with other musicians in chamber music and orchestras’ (LHPP: 240). Hegel agrees with Humboldt (Hardimon 1994: 188), and Hegel’s idea of a well-ordered society may be described in the same terms: ‘Magnificent instrumentalists as we may be, we can only achieve the consummation of our skills alongside others in the social orchestra’ (Knowles 2002: 56). See also Inwood (1992: 107, 313). On Rawls’s orchestra analogy, see Love (2003).

providing labour incentives (Van Parijs 2003: 203-4; CP: 257). More fundamentally, the role of inequalities in distributive shares is to

attract people to positions where they are most needed from a social point of view, to cover the costs of acquiring skills and educating abilities, to encourage them to accept the burdens of particular responsibilities, and to do all this in a way consistent with free choice of occupation and fair equality of opportunity (JF: 78).

This does not assume that persons deserve their place in the distribution of natural endowments and therefore whatever market rewards they can achieve with them. Rather, in viewing the *distribution* of those natural endowments as a common asset, a ‘basic structure satisfying the difference principle rewards people . . . for training and educating their endowments, and for putting them to work so as to contribute to others’ good as well as their own’ (JF: 75). The difference principle is a principle of fraternity, of ‘not wanting to have greater advantage unless this is for the benefit of others who are less well off’ (CP: 167). Certainly Rawls’s conception of a social union of social unions appeals to Adam Smith’s invisible hand (TJ: 57, 525 n.4/49, 460 n.4), but so, we should remember, does Hegel’s (LHMP: 346). Such a view is consistent with the Kantian idea of treating persons as ends and never as means *only*, since when citizens precommit themselves to principles to regulate the operation of the market they freely consent to, and publicly direct, the operation of the invisible hand.

Is Rawls’s position an example of ‘an infantile unwillingness to countenance a measure of self-denial as a way of dealing with the inevitable difficulties of social existence’ (SFE: 135)? It certainly doesn’t reject a *measure* of self-denial on the part of citizens, who agree to regulate their desires. But to maintain our reasonable faith in the possibility of a well-ordered society, self-denial should not become self-abnegation: the most stable conception of justice is one that is ‘perspicuous to our reason, congruent with our good, and rooted not in abnegation but in affirmation of the self’ (TJ: 499/436). It is only because Cohen regards all confinement according to rules as coercive that he rejects a political-institutional approach to justice. In one sense, Rawls admits, justice as fairness, unlike Marx’s view of justice, is coercive, because for Marx ‘no one can be required to benefit himself only in ways that contribute to others’ well-being’ (LHPP: 368). We might think that ideally persons would serve one another (TJ: 151/131): Rawls did once describe incentives as ‘concessions to human nature’. But in describing how incentives ‘work to draw out better efforts’, already there is a suggestion of the ideal of reconciliation to our social world in the idea that acceptance of these inequalities by members of society is ‘acceptance of the relations in which they actually stand’ (CP:

55). Rawls's difference principle is not the same as the maximin choice rule, but the idea of making the worst off as well off as they can be takes us back to Rawls's idea of realistic utopianism and reconciliation to our social world. Adopting a utopian conception of equality, the family, for example, will appear simply as an obstacle. Like Plato, we might be led to think about whether the family should be abolished (Munoz-Dardé 1999). But after asking 'Is the family to be abolished then?', Rawls suggests that public recognition of the difference principle 'seems more likely than its rivals to transform our perspective on the social world and to reconcile us to the dispositions of the natural order and the conditions of human life' (TJ: 512/448; see also JF: 76).

In assessing Cohen's utopian socialism, one could point out from a traditional Marxian perspective the arbitrariness of moralising critique that ignores the constraints of existing institutions. Rather than blaming the lack of self-denial of the best off, what about the subservience of the least well off, why not 'If you're an egalitarian, how come you're so poor?' (Sensat 2003: 340). But from the reformist perspective I am concerned to defend, the real problem with Cohen's position is that in purporting to offer an internal critique of Rawls, Cohen's absents his own normative standpoint from assessment according to an interpersonal test. The values of democratic community should apply to procedures of political justification and not just be applied to their object domain (Bertram 1997). One might ask: If you're an egalitarian who believes that occupational choice is answerable for at the bar of justice, how come you're a philosopher? (see Scialabba 2001). Is 'seeker after truth' a privileged position that excuses one from working to benefit the least well-off (subject of course to an agent-centred prerogative)? If not, then one would need an argument for the intrinsic practical value of the truth about justice.

Rawls's strategy of working with the grain of natural desires is based on a vision of justice as reciprocity, as opposed to justice as a utilitarian higher-order executive decision. It is to the latter that Cohen's approach would condemn us – justice as the engineering solutions of socialist economists – while we await a revolution in morals. Reciprocity is situated between altruistic impartiality and self-interested mutual advantage (CP: 77; PL: 16-17; Gibbard 1991), and is therefore between community and bargaining in Cohen's sense. But this is because Rawls rejects the idea that a Cohenite notion of community is true community. In his undergraduate conception of community that continues to resonate in his later work, Rawls distinguishes between natural relations to objects through which one satisfies desires and personal relations through which persons establish moral relations with an 'other'. Cohen argues for natural relations to fundamental principles of justice, not an ideal of personal relations, making the same mistake for which we will see that Rawls criticises utilitarianism, conflating

impartiality and impersonality. Comprehensive justification in justificatory community is not a reconciliation of conflicting interests but acquittal at the bar of justice that places one beyond the claims of others.

We have seen that Cohen reduces Rawls's concern to preserve a nonpublic space for the development of an autonomous social conception of the good – in accordance with reasonable pluralism, and on the basis of which principles of justice may be autonomously endorsed – to the level of a pragmatic compromise to acquisitive capitalist self-interest. This is a bleak, and indeed unrealistic, view of contemporary liberal societies if it implies that there are no social spaces beyond the capitalist market in which political principles of justice may find a social foothold. On the other hand, whilst Scheffler moves beyond the idea that the division of moral labour is a way of externalizing the demands of morality, and recognises Rawls's social holism, he does not adequately explore the more fundamental role of social institutions for Rawls in educating persons to a conception of themselves as free and equal citizens. Respect for moral pluralism is not a *foundational* value in Rawls's approach, but rather a social space of autonomy is recognised *within* Rawls's theory insofar as this is a precondition for persons' autonomously endorsing political principles of justice.

5. Rescuing Constructivism from Platonism and Positivism

The previous chapter showed how for Rawls institutions are not vehicles for the instrumental realization of independently derived values. In Rawls's political constructivism, a conception of justice is constructed in accordance with a union of practical reason and ideals of society and the person, and assessed from within the basic or background institutions of the basic structure, such reflective endorsement being the basis for reforming the basic structure as an ongoing practice of social cooperation. Compared to Cohen's understanding of constructivism, this may appear outlandish, with little connection to what Rawls is standardly thought of as doing with the original position in *TJ*. But 'constructivism' is an infelicitous term for what Cohen seeks to oppose. 'Contractarianism' would have been closer to the mark, since Cohen claims his 'critique of Rawlsian constructivism also applies, *mutatis mutandis*, to Scanlonian contractarianism, to Gauthier's contractarianism, and to Ideal Observer theory' (RJE: 275). Cohen's methodological critique is, moreover, driven by concrete political issues lying at some remove from Rawls's methodological concerns, constructivism's acceptance of incentives standing as one example of its pragmatic accommodation to existing facts.

Having said this, employing Cohen's critique as a foil for interpreting Rawls is a fruitful way of drawing attention to fundamental issues of philosophical methodology. Beginning from an 'Aristotelian' way of thinking about justice, Rawls believes that a shared sense of justice makes a just democratic society and that philosophical justification should seek to give structure to common sense judgements about justice in a manner that respects persons' sense of justice. I have argued that in a Hegelian fashion, Rawls develops the former commitment into a conception of a well-ordered society in which a Rousseauian general will is realized in the institutions of society. This chapter extends my analysis of the latter commitment, the way in which the role of a constructivist procedure within reflective equilibrium is to provide a structure that incorporates rather than transcending its object and is to this extent similar to Hegel's dialectics. This results, I will argue, in an internal relation between principles chosen in the original position and facts that are constitutive of the basic structure as an ongoing practice of social cooperation. Both social holism and dialectical thinking are aspects of the Hegelian 'bullshit' that 'non-bullshit' analytical Marxism sought to dispense with (KMTH: xxv). In first comparing Cohen's position with that of Moore, I seek to show how the presuppositions that formed analytic philosophy in its reaction against idealism

continue to inform Cohen's work. A corollary of my defence of Rawls will be that this analytic reaction was an overreaction, offering implicit support for efforts in postanalytic philosophy to come to terms with idealist conceptions.

5.1 Platonic Atomism and Quinean Holism

In Rawls's constructivism, the 'very content of the first principles of justice . . . is determined in part by the practical task of political philosophy' (CP: 330). Cohen argues, by contrast, that practical considerations are irrelevant to theoretical philosophical analysis of fundamental principles of justice. Constructivism for Cohen is the view that principles gain their normative grounds by being the output of a justifiable selection procedure. Cohen's focus is constructivism with respect to fundamental principles of justice and more specifically Rawlsians who are said to

believe that the correct answer to the question 'What is justice?' is identical to the answer that specially designed choosers, the denizens of the Rawlsian original position, would give to the question 'What general rules of regulation for society would you choose, in your particular condition of knowledge and ignorance? (RJE: 277).

This is taken to lead to a double misidentification. First, in making the choice of rules of regulation subject to factual constraints, these rules are made to depend upon considerations regarding fundamental principles generally and not upon considerations that serve principles of justice in particular. Cohen conceives of rules of regulation in teleological terms: they are a 'device for having certain effects' (Nozick, quoted in RJE: 327). Second, fundamental principles of justice are made to depend upon considerations other than considerations of justice. On Cohen's alternative view, justice is seen as one fundamental principle among others. Rules of regulation can therefore seek to serve justice in particular, while recognising that there are other fundamental values. For Cohen, constructivists misidentify fundamental principles of justice because they take account of other values relevant only to optimal principles of regulation. 'It follows that any procedure that generates the right set of principles to regulate society fails thereby to identify a set of fundamental principles of justice, by virtue of its very success in the former, distinct, exercise' (RJE: 283).

Cohen's approach therefore involves identifying fact-independent fundamental principles of justice and choosing rules of regulation to realize as much justice as possible in the circumstances, subject to intuitive trade-offs against other fundamental principles. This means that Cohen affirms the intuitionism, or 'radical pluralism' (RJE:

4), that Rawls opposes, according to which there exist no ‘higher-order constructive criteria’ for balancing first principles against one another (TJ: 34/30). While such radical pluralism is opposed to hedonistic utilitarianism, it is exemplified by the ideal utilitarianism of Moore. Before turning to Cohen’s argument against constructivism, it will help to clarify what is at stake in this debate by contrasting Cohen’s Moorean intuitionism with Rawls’s Quinean holism.

For Moore, the ‘fundamental principles of Ethics must be self-evident’, and all moral laws are ‘merely statements that certain kinds of actions will have good effects’ (Moore 1993 [1903]: 193, 196). Moore’s (1993 [1922]) argument for a conception of intrinsic value is based on what, following Peter Hylton (1990), Rawls calls, a ‘Platonic atomism’ that sees moral and other concepts as independent entities grasped by the mind (CP: 344, 511). This positive argument for Platonic atomism works in tandem with a negative argument against the ‘naturalistic fallacy’ committed by both hedonistic utilitarianism and Kantian metaphysical ethics. In asking whether a thing has intrinsic value, we are asking about the intrinsic nature of the things in question, and this intrinsic nature consists of a thing’s non-relational properties, properties that it would have in any possible world. The pluralism of Moore’s view derives from his principle of ‘organic unities’. While intrinsic value is usually a property of organic wholes, and the value of a whole may be greater than the sum of its parts, there are no rules for determining how the value of each part contributes to the value of the whole. We can only determine the intrinsic value of a thing by an intuitionistic ‘method of isolation’, considering ‘what things are such that, if they existed *by themselves*, in absolute isolation, we should yet judge their existence to be good’ (Moore 1993 [1903]: 236).

The method of analysis depends upon being able to distinguish between analytic propositions that are true by virtue of the meaning of words and synthetic propositions that are dependent upon experience, and therefore facts. Rawls follows Quine in rejecting this distinction. According to Rawls’s Quinean concept-theory holism, no sharp distinction can be drawn between the concept of justice and conceptions of justice, with the result that we can only seek to articulate the concept of justice through a particular conception (Hardimon 1994: 97-8). ‘The merit of any definition depends upon the soundness of the theory that results’ (TJ: 130/112-13). Just institutions may be considered as abstract objects – ‘possible form of conduct expressed by a system of rules’ – but as such an institution is ‘just or unjust in the sense that any realization of it would be just or unjust’ (TJ: 55/48). In a Hegelian sense (LHMP: 350), this involves the ‘realization in the thought and conduct of certain persons at a certain time and place of the actions specified by these rules’ (TJ: 55/48).

When Rawls says ‘[t]he principles of justice define an appropriate path between dogmatism and intolerance on the one side, and a reductionism which regards religion and morality as mere preferences on the other’ (TJ: 243/214), one should read this against the background of Quine’s rejection of two dogmas of empiricism. The first of these connected dogmas is the distinction between analytic and synthetic propositions, which Quine argues is a matter of degree rather than a categorical distinction. The second is a verification theory of meaning according to which ‘the meaning of a statement is the method of empirically confirming or infirming it’, where an analytic statement is a ‘limiting case which is confirmed no matter what’ (Quine 2001 [1951]: 457). As this suggests, the two dogmas support one another: the dependence of synthetic statements on verification by empirical experience supports the view that statements that cannot be so confirmed are analytic.

While Quine’s holism, and its rejection of any dualism of form and content, is reminiscent of Hegel, it is of course true that ‘Quine’s emphasis on natural science, in particular, is a fundamental point of disanalogy’ (Hylton: 484 n.58). However, this is not a naturalism that Rawls fully endorses. Rawls distinguishes his approach from both Cartesianism and naturalism (TJ: 578/506). Cartesianism proceeds by deduction from self-evidently true first principles, while naturalism conceives of moral justification on the model of scientific justification and seeks to define moral concepts in terms of other concepts supposed to be non-moral. Against Cartesianism, Rawls argues that there are no first principles that can carry the burden of the justification of morality. Against naturalism, on the other hand, he argues that there is no theory of meaning that can unproblematically distinguish between moral and non-moral terms. Emphasising the political motivations that drive Rawls’s constructivism, we can say that it seeks to avoid both Platonism and positivism, the former being the mirror image of the later. This concern goes back to Rawls’s doctoral thesis where what I have called Platonism is an example of what Rawls there labels ‘authoritarianism’, and that he seeks to avoid as much as a positivism, which, whether in the form of ‘emotivism, social structuralism or pure theory of law’, is based on the belief that moral principles cannot be ‘validated or invalidated by reason’ (SGEK: 3-4).

The difference between Cohen and Rawls may be illuminated with reference to Korsgaard’s (1996c) differentiation of two distinctions in goodness. On Cohen’s Moorean conception, rules of regulation are instrumental to the realization of the intrinsic good of justice. But, as Korsgaard argues, this position runs together two separable distinctions. The converse of intrinsic goodness – goodness in itself – is not instrumental goodness but extrinsic goodness – goodness that is derived from another source. This question of the source of goodness is different from the question of whether

something is valued for its own sake or for the sake of something else. We now see more clearly the underlying differences between Cohen's and Rawls's conceptions of justificatory community. A well-ordered society for Rawls is something that is *extrinsically* good. Its goodness comes from persons' sense of justice, the fact that they value such a society for the right reasons. However, given favourable circumstances that allow for full compliance, a well-ordered society can be valued as an *unconditionally* good end, one that is not dependent upon persons' particular ends.

The problem with Moore's practice of isolating the intrinsic value of organic unities by intuition is that it 'conceals the fact that an organic unity has value in virtue of its structure, of the internal relations of its parts' (Korsgaard 1996a: 112 n.23). Moore's view of ethics derives from his rejection of British Idealism and idealists' use of a doctrine of internal relations (Moore 2000 [1903], 2000 [1919]). But lest it be thought that this is simply a necessary corrective to the speculative excesses of idealism, it is important to note that Platonic atomism is opposed not simply to the speculative metaphysics of British Idealism, but 'quite generally opposed to all forms of Idealism, including, as Russell and Moore held, Kantianism' (Hylton 1993: 449). Moore rejects Kant's 'Copernican Revolution', the shift from the analysis of objects to the analysis of the necessary conditions of our experience of objects, or transcendental analysis, such conditions being neither empirical in the sense of being psychological or physiological, nor simply logical. For Moore, 'it is plain that the only connection which can possibly hold between being true and being thought in a certain way, is that the latter should be a *criterion* or test of the former' (Moore 1993 [1903]: 183). By contrast, Kant's analysis of how the mind actively constitutes the knowable world opens the door to understanding objectivity and subjectivity as interrelated, and the idea of a viewpoint that is based in practice but can claim objectivity. Both Russell and Moore on the one hand and post-Kantians on the other accept Hegel's critique of Kant, but they respond in opposite ways. While Platonic atomism disengages the mind from the constitution of knowledge, the Hegelian strategy is to locate the mind in a social and historical context in order to vindicate knowledge against scepticism regarding its objectivity.

Viewing Rawls's original position from this perspective we may note that its justification depends upon how it 'fits in with and organizes our considered judgements in reflective equilibrium' (TJ: 579/507). It cannot be assessed alone, but only in terms of the theory of which it is a part. Two aspects of Rawls's position have been regarded as constructivist, of which Cohen considers only the first. The first is the construction of a conception of justice through the constructive procedure of the original position, while the second is what Dworkin has described as the constructive process of reflective equilibrium through which the original position itself is justified (Krasnoff 1999).

Rawls's interpretative method is different to Dworkin's, but, notwithstanding the scepticism of those such as Barry (1989: 276), Rawls does employ the original position within such a hermeneutical methodology. Rawls employs a constructivist procedure *within* a philosophical division of justificatory labour between moral conditions and psychology and social theory. In 'Kantian Constructivism in Moral Theory', the enterprise of moral theory at issue is a matter of how different substantive moral conceptions arrange notions of the right, the good and moral worth to form different moral structures (CP: 286). With his political turn, Rawls argues that with hindsight this should have been called 'Kantian Constructivism in Political Philosophy' (CP: 389 n.2). In *PL*, in specifying how constructivism's ideas are drawn from the domain of the political, Kantian constructivism becomes political constructivism. In turning to Cohen's critique of Rawls, we will need to bear in mind that 'reflective equilibrium works *through* the original position' (Rawls, quoted in Freeman 2009).

5.2 Facts and Principles

Cohen's critique of constructivism concerns the relationship between facts and normative principles. Facts are defined negatively as any truth other than a principle, while a principle is defined as a general directive that tells an agent what they ought or ought not to do. The focus of the argument is on the logical structure of an individual's beliefs. Cohen's 'clarity of mind requirement' specifies that his thesis applies to anyone's principles

so long as she has a clear grasp both of what her principles are and of why she holds them (where 'grasping why she holds them' is short for 'knowing what she thinks are the grounds of the principles' rather than for 'what causes her to hold them'). It also characterizes (under an appropriate reformulation) whatever (if anything) constitutes the *correct* set of principles (RJE: 233).

Cohen's thesis is conditional on there being persons who affirm principles on the basis of facts. Constructivists in general, and Rawls in particular, are held to adopt this position and as evidence Cohen cites Rawls's claim that 'Conceptions of justice must be justified by the conditions of our life as we know it or not at all' (TJ: 454/398). More specifically, the belief that Cohen says constructivist hold is that our beliefs about normative principles should reflect, or respond to, truths about matters of fact, where by this is meant that they should 'include matters of fact among the grounds for affirming

them' (RJE: 231). If facts enter into the grounds of belief in this way then the principles endorsed are fact-sensitive.

The intention of Cohen's argument is to show that not all principles are fact sensitive. His thesis is that '*a principle can reflect or respond to a fact only because it is also a response to a principle that is not a response to a fact . . . [P]rinciples that reflect facts must, in order to reflect facts, reflect principles that don't reflect facts*' (RJE: 232). Cohen's argument has three premises, and the third premise rests on three subsidiary arguments:

1. '[W]henever a fact *F* confers support on a principle *P*, there is an explanation *why F* supports *P*, an explanation of how, that is, *F* represents a reason to endorse *P*' (RJE: 236);
2. '[T]he explanation whose existence is affirmed by the first premise invokes or implies a more ultimate principle, commitment to which would survive denial of *F*, a more ultimate principle that explains *why F* supports *P*' (RJE: 236);
3. *If* a person holds a principle on the basis of a fact, then if anyone who does so is asked what more ultimate principle explains why that fact grounds that principle, and we proceed reiteratively as many times as may be required, then this process will come to rest with a principle that reflects no fact and will not proceed indefinitely. This is because:
 - a) No plausible counter-example can be constructed that begins to even suggest such a process could proceed indefinitely;
 - b) An infinite sequence assumes an implausible infinite nesting of principles;
 - c) An infinite sequence would contradict the clarity of mind requirement that she who affirms a principle 'has a clear grasp of what her principles are and of why she holds them' (RJE: 237).

5.3 Beyond Facts and Principles

5.3.1 Objection to the first premise

Cohen's first premise is part of a more general claim, unargued but taken to be self-evidently true, that where there is a grounding relation of fact-sensitivity there is always an explanation of why any ground grounds what it grounds. However, the nature of the grounding relation, the relation through which *F* confers support on, or provides a reason to endorse, *P* is ambiguous.⁸⁰ Cohen employs a number of locutions, but the ambiguity can be brought out by considering the difference between a principle *reflecting a fact* and a principle *responding to a fact or facts*. For a principle to reflect a fact suggests that a single factual statement is incorporated into the justification of a principle such that when the compound is analysed the linguistic meaning of the principle will be separable from the factual component. As the rational basis of the principle, it would be a requirement of theoretical reason to trace back the deductive argument on the basis of which a fact entered as a reason to support a principle. It is the idea of reflecting a fact by incorporating it into the grounds of a principle that makes plausible Cohen's formulations where the relation *R* of a fact *F* to a principle *P* is one where *F* 'grounds' or 'justifies' *P*, or where we are committed to *P* 'because of' or 'on the basis of' *F*.

On the other hand, the responsiveness of a principle to a fact suggests a matter of degree.⁸¹ Take as a candidate 'fact' that individuals require material incentives to work at their full productive potential. If we seek to respond to that fact we might place it in a larger context of a whole web of facts about persons' familial, social, cultural, religious and political relationships and commitments. More abstractly still, we might place these within an account of the operation of human psychology within social institutions. The original fact may then be regarded as a brute fact, brute, that is, relative to these higher level descriptions (Anscombe 1958a). We could also say that a brute fact has become an institutional fact as a result of identifying the presupposed context that constitutes the fact as a fact. Then we could distinguish between regulative rules that 'regulate antecedently or independently existing forms of behavior' and constitutive rules that 'create or define new forms of behavior' (Searle 1969: 33). The public system of rules of the basic structure includes constitutive rules of social cooperation in this sense (TJ: 56/49).⁸²

⁸⁰ The following analysis is indebted at a number of points to Guay (2006).

⁸¹ As O'Neill (2003a: 363 n.3) notes, etymologically 'facts are originally *facta* – done or made by agents rather than *given* to them, as *data* are'.

⁸² On Searle in relation to Rawls, see Karp (2009).

We can understand the idea of a principle responding to a fact, or facts, as a form of deliberative or practical sensitivity to facts in contrast to a rational or theoretical sensitivity (Guay 2006). It is in this sense that we should understand the idea that principles are chosen in the original position *in the light of* facts. It is certainly true that in *TJ* the facts allowed through the veil of ignorance can appear ad hoc, but Rawls is consistent in referring to them as *general facts*. I argued in the previous chapter that in forming the background to our lives, and therefore our social world, the basic structure is akin to a hermeneutical horizon of understanding within which moral reflection takes place. In constructing principles that respond appropriately to these facts, Rawls's original position incorporates constraints on the choice of principles of justice – or, we could say, categories – that are presuppositions of a practice of social cooperation.⁸³

The next step in understanding Rawls's view is to connect his incorporation into the original position of general facts about our social world with his description of justice as fairness as 'a theory of the moral sentiments . . . setting out the principles governing . . . our sense of justice' (TJ: 51/44). Justice as fairness is the hypothesis that the principles which would be chosen in the original position are identical with those that match our considered judgements in reflective equilibrium and therefore describe our sense of justice (TJ: 48/42). In endorsing such a conception in reflective equilibrium, we 'understand how the moral sentiments can be regulative in our life and have the role attributed to them by the formal conditions on moral principles' in the original position (TJ: 478/418).

The constitutive rules of justice for the reform of the basic structure of society will therefore be shown to be those that are regulative of our sense of justice. These principles must therefore cohere with our considered judgements about justice, such as that slavery is unjust, as fixed points on the basis of which we undertake a Socratic process of seeking self-knowledge about our sense of justice. An example of the first kind of fact with which constructivism is concerned is that to illustrate the fact that slavery is unjust we appeal to the fact that it allows some persons to have property rights in other persons.

With regard to the first kind of relevant facts, a constructivist procedure is framed to yield the principles and criteria that specify which facts about actions, institutions, persons, and the social world generally, are relevant to political deliberations . . . Apart from a reasonable moral or political conception, facts are simply facts. What is wanted is a framework of

⁸³ See Miller's (2008: 33-34) criticism of Cohen's narrow understanding of the grounding relation as one where 'facts ground principles by virtue of being premises in a relationship of logical entailment' and alternative conception of presuppositional grounding where rather than A entailing B, 'A's being true is a necessary condition of B's being true'.

reasoning within which to identify the facts that are relevant from the appropriate point of view and to determine their weight as reasons (PL: 122).

By giving coherence to our considered judgements, justice as fairness aspires to allow us to make sense of them by seeing them in an appropriate light. Cohen argues that ‘facts cast normative light only by reflecting the light that first principles shine on them’ (RJE: 267). But for Rawls facts about the operation of persons’ sense of justice are not value free. Facts cited in support of our considered judgements, facts, that is, that are expressive of our sense of justice, cast normative light.

We cannot observe our sense of justice from an impersonal standpoint, for to seek self-knowledge in understanding its operation from an abstract standpoint is to engage in an act of reflection through which our sense of justice will be altered. We only understand our sense of justice in the process of practical deliberation. The mirror that justice as fairness holds up to society cannot for that reason simply capture a reflection. The process of reflective equilibrium is one in which the original position is an ‘idealizing mirror’ (Ron 2006), a mirror that shows us the principles of justice that are regulative of our sense of justice under ideal but realizable conditions. As in Rawls’s description of Kant’s categorical imperative procedure, a conception of persons as free and equal is ‘mirrored in the procedure’ (LHMP: 240). Since Rawls takes persons to be capable of reasonable deliberation regarding this idealized reflection, he takes them to be motivated to endorse this idea when they can regard it as in fact reflecting their sense of justice, and therefore their status as free and equal citizens. ‘In light of the theory of justice we understand how the moral sentiments can be regulative in our life’, since such philosophical reflection ‘bring[s] to light’ the source of morality in our person (TJ: 478/418). Then, however, persons’ sense of justice will cast a new – brighter, one could say – normative light, one that we can again reflect on from an abstract standpoint in an endless reflexive and reiterative process. In this back and forth reflection we may think there is an external light source, but that would represent an estrangement from the source of principles in our sense of justice. W.D. Ross complained that to make value relational would make it the case that ‘value would seem always to be borrowed, and never owned; value would shine by a reflected glory having no original source’ (quoted in Korsgaard 1996: 256). But for Rawls we do not reach a point where our considered judgements are true in the light of an independent order of values; rather in ‘using our reason to describe itself . . . the struggle for reflective equilibrium continues indefinitely’ (PL: 97).

5.3.2 Objection to the second premise

On the basis of this understanding of the original position as an idealizing mirror held up within our social world, we can turn to Cohen's second premise (that a grounding relation implies a more ultimate principle that explains why a fact grounds a principle). Cohen defends this premise by challenging objectors to provide an alternative candidate for the place occupied in his schema by fact-insensitive normative principles. He considers first the idea that the original position is a more ultimate principle, but one that is methodological not normative, that 'tells you how to choose principles that tell you what to do', not directly what to do (RJE: 240). But, Cohen says, 'when the original position machine selects *P* in the light of a set of factual truths, that is because it would . . . select a fact-free normative principle *P1* when those factual truths are suspended' (RJE: 240).

Cohen's argument would have some force if the original position were akin to the categorical imperative. The criticism would then be that the original position is dependent on a conditional rather than an unconditional end. But the higher-level categorical imperative procedure identified in such a way would not be a higher-level *substantive* principle. It would also not be fact-independent – as the moral law that governs our practical reasoning as autonomous moral beings the categorical imperative retains a connection to facts about the kind of beings that we are. Alternatively, Cohen might mean that conditional principles can be made to count as fact-independent by the addition of a hypothetical 'if'. Cohen, for example, gives as an example of a fact-independent principle 'If a being is liable to pain, you ought not to cause it pain' (Miller 2008: 35 n.8). Pogge (2008) refers to this as making externally fact-sensitive principles internally fact-sensitive. But if this is the case then his argument risks becoming trivial.

The more fundamental issue, however, is that the original position is not justified independently of our considered judgements about justice: we design the original position so as to get the desired results – or rather so that its results accord with our considered judgements – as well as seeing results as desirable because they were chosen in the original position. Rawls's original position can hope to justify substantive principles because it claims to embody conditions on the choice of principles that we *do in fact accept*, in the sense that they are latent in common sense and we would recognise them on due reflection (TJ: 21, 587/19, 514). It makes no sense to ask what principles would be chosen within the original position under different factual conditions as the original position is set up by us, within our social world, as a device of reflection and self-clarification.

Cohen's second objection is that 'we have to reckon not only with the principles justified by the original position procedure, but also with the principles that justify that procedure. Procedure is not ultimate: as Rawls says [PL: 103], not everything is constructed' (RJE: 241). This echoes a familiar line of criticism that Rawls must be relying upon a foundational principle of equal concern and respect to justify the use of the original position (Dworkin 1989). But for Rawls respect for persons is fundamentally shown by 'treating them in ways that they can see to be justified' (TJ: 586/513). Treating persons as ends and not merely means in matters to justice is to treat them in a way that is publicly justifiable to common human reason (LHMP: 192). We begin to cash this out by assessing how we can give social realization to the relationship of freedom and equality constituted within the structuring conditions of the justificatory device of the original position.

Rawls's point is that it would not even strike us as reflectively plausible to think about justice in terms of principles chosen in the original position unless we *already* in some sense believed that persons were free and equal. Principles of justice must be able to explain, for example, why we take slavery to be unjust. This considered judgement is not one isolated 'intuition', but is connected with facts like tyranny is unjust, exploitation is unjust and religious persecution is unjust. 'Political constructivism does not look for something for the reasonableness of the statement that slavery is unjust to consist in, as if the reasonableness of it needed some kind of grounding' (PL: 124). That facts about our considered judgements can be connected by principles constructed by a procedure that incorporates the requirements of practical reason together with conceptions of the person and society 'is not a fact behind all the separate facts; it is simply the fact of these connections now open to view and expressed by the principles free and equal persons would accept when suitably represented' (PL: 124-25). This, the second kind of facts with which constructivism is concerned, involves facts about a political conception of justice. When we work out a political conception from the fundamental ideas of a well-ordered society, it 'is a possibility of construction, implicit in the family of conceptions and principles of practical reasoning that are the basis of the construction, that slavery is unjust' (PL: 123). Here 'basis' should be understood not as a foundation but, in the manner of Neurath's boat, as the social world upon which we are afloat when reconstructing its basic structure of social cooperation. Cohen is right to suggest that procedure depends upon substance, but wrong to think that the idea that persons are to be regarded as free and equal need rely upon a fact-insensitive normative principle outside the procedure of construction rather than a principle implicit in our substantive commitments.

5.3.3 Objection to the third premise

Turning to Cohen's third premise (that we can proceed reiteratively but not indefinitely in asking what more ultimate principles are implied by the grounding relation between a fact and a principle), Cohen's qualification ('if a person holds a fact on the basis of a principle') is important. His focus is persons' first-personal awareness of the rational structure of their beliefs, rather than the causal or psychological question of what causes persons to believe things. Cohen recognises that his argument depends upon a foundationalist view of justification that would be rejected by holistic or coherentist views. But, citing Jerrold Katz (1998), Cohen argues 'truths that determine the impact that our statements have upon one another, within the full set of our statements, cannot themselves be treated holistically, on pain of infinite regress' (RJE: 242).⁸⁴ This is true to the extent that not all beliefs within a holistic system of justification are, or can be, on the same level. If that were so our belief systems would have no structure. But is it necessary that the structure of our beliefs be determined by an unrevisable truth outside our system of beliefs? Whether there are such facts is precisely the point at issue. It is not, as Katz argues, that holistic justification saws off the branch on which one is sitting, but rather that in a bootstrapping virtuous spiral, common sense is systematised according to standards that common sense implicitly accepts.

Cohen's clarity of mind requirement is fundamental to his argument. It is what makes it, as far as he is concerned, a philosophical thesis immune from psychological facts. But scepticism will enter if the clarity of mind requirement looks so demanding that it implies that constructivists are *really* committed to a comprehensive egalitarian doctrine akin to a religious worldview. Cohen conceives of reflection about justice on the model of individual reflection on the demands of theoretical rationality. But Rawls's point is that such reflection is always too indeterminate to structure a practice of social cooperation. What defines the reasonable is that it is public in a way that the rational is not. Rawls's argument depends on the claim that the domain of the political, like the social practice of a well played game, has its own values, and that the fact that we cannot realize within it what we may, as a holder of a comprehensive doctrine, take to be the whole truth need not be seen as a compromise but rather as realizing the values that are integral to political cooperation.⁸⁵ Rawls says

⁸⁴ Katz's Platonic realist rationalism makes an interesting comparison with Cohen's view. See Balaguer (2003); Lackey (2003).

⁸⁵ Cohen himself recognises that games are a counter-example to Murphy's argument for monism (RJE: 397-98).

that we think political values have some further backing does not mean we do not accept those values or affirm the conditions of honouring public reason, any more than our accepting the axioms of geometry means that we do not accept the theorems. Moreover, we may accept the axioms as much because of the theorems they lead to as the other way round (PL: 242).

We must distinguish between the order of deduction and the order of support (PL: 242 n.31), that is, between the deductive and inductive aspects of Rawls's argument and how it is the inductive aspect, the process of reflective equilibrium, upon which full justification depends.

We have seen how the ambiguity in Cohen's first premise allows him to present the normative principles constructed by constructivism, which respond to, or are constructed in the light of, facts as incorporating and reflecting unchanged a fact into those principles' rational basis of justification. Cohen's second premise then makes the unwarranted assumption that such fact-sensitive principles imply more ultimate normative principles. This would be untrue of a principle such as the categorical imperative, which, although it exists independently of the phenomenal world, is not a normative principle in Cohen's sense. It is even more untrue of the original position, which does not depend upon a more ultimate principle of respect for persons, but models respect for persons within its construction and is ultimately justified by the reflective endorsement that serves to incorporate this relationship of respect into the basic structure of society. Contrary to Cohen's third premise, this process *does* proceed indefinitely because reflective equilibrium proceeds indefinitely. However, a final point needs to be addressed. Cohen's second reply to the holism objection is to argue that even if some form of global holism were granted there will be some statements that are virtually independent of others and that is enough for his argument to go through. Indeed, Cohen rightly says that his substantive dispute with constructivism concerns a far more local sensitivity of principles to facts. I turn to this in the next section.

5.4 Principles of Justice, Rules of Regulation and Principled Rules of Regulative Justice

Cohen says that for constructivists to press a holism objection ‘they have to claim that the local context under discussion, the context of *facts and principles they support*, is *itself* holistic’. Moreover, they must ‘provide an illustration of how a change in one’s view of a *principle* can alter one’s belief about the *sort* of fact that they think supports principles’ (RJE: 243). I have already indicated in Chapter 2 the way in which moving from regarding principles of justice as utilitarian principles in accordance with a liberalism of the good, to a liberalism of freedom that seeks to describe our sense of justice in reflective equilibrium meets the second challenge. I take up the first challenge in this section by arguing for a local holism consisting in an internal relation between principles of justice and the general facts, or circumstances, of our social world.

The distinction between external and internal relations can be put in the following terms:

If one item, x , stands in some relation, R , to another item, y , but neither its identity nor its nature depends upon this being the case, x is *externally* related to y . If x could not be the same item, or an item of the same kind, without standing in relation R to y , the relation is *internal* (Bogen 1995).

Rawls’s position is that principles of justice, x , stand in a relation of reflective equilibrium, R , to our social world, y , and the nature of this relationship as one *within* a process of reflective equilibrium is such that x and y are internally related. If this is correct, then Cohen’s claim that his critique of constructivism is neutral with respect to the site of justice will simply beg the question. It will only be shown to be neutral if Cohen can successfully carry through the argument that principles of justice are separable from all facts about the world and show that the facts that Rawls does invoke are unrelated to his focus on the basic structure of our social world as the first subject of justice.

The general facts that are relevant for justifying principles of justice are facts that are constitutive of the possibility of our social world taking the form of a practice of ongoing social cooperation between free and equal persons. Principles of justice would cease to be principles of *justice* if they could not be understood as capable of structuring a practice of social cooperation; they would remain a mere logical proof and would not appear as justified to common human reason. Justice has a dual reference, referring to an implicit possibility of construction and a constructed conception of justice, an existing social practice and principles that morally require the reform of that social practice.

The idea of reflective equilibrium originates in Rawls's doctoral thesis. There Rawls asks the question 'Do principles exist which are so well grounded that we do not have to appeal to *another* principle in order to justify them?' (SGEK: 271). His answer is posed in terms of principles for assessing the moral worth of character, but it is readily transferable to his later concern with principles of justice for the basic structure.

It is true that the appeal to *further* principles must stop; and in this sense we cannot provide a justification for the principle at which the argument ends. But it is a fallacy to believe that we cannot support this principle by a different kind of consideration . . . [T]he ideal of rational action . . . represents that ideal or norm of action which the unhampered and uncoerced and collective sense of right has decided upon as just and fair. Thus while we cannot bring the ideal under a higher principle we can show that the principle appealed to is not arbitrary or capricious, but reflects the deepest convictions of reasonable people. In this way we can justify the stopping point, not by reference to another norm, but by the kind of allegiance gained by seeing its rightness after long criticism and reflection, and weighing it with alternatives. We may say of it that it is willingly adopted, of itself, by the free collective sense of right (SGEK: 271-72).

The original position represents persons as free and equal, excluding specific facts but excluding no general facts about the circumstances of our social world. The formal constraints of the concept of right eventuate in a principled choice of constitutive rules realized within the basic structure of society, understood as a public and self-reproducing practice of social cooperation. The full justification of these principled rules depends upon them being reflectively endorsed such that we recognise them as our regulative sense of justice within which we desire to pursue our conceptions of the good. Rawls is not concerned with identifying principles of justice in Cohen's sense, but nor is he interested in deriving rules of regulation. Instead, we can see Rawls as seeking *principled rules of regulative justice*. This argument is circular, but not viciously so, and because principles of justice are defined in the light of conditions that are regulative of our sense of justice, and our sense of justice is generated within the institutions formed by principles of justice, the two cannot be separated. We could even describe Rawls's constructivism as a transcendental hermeneutics of our social world. Publicity and stability are categories appropriate to a deliberative procedure for identifying which facts are relevant when it comes to determining principles for the reform of the basic structure of society.

The original position brings to consciousness the principles that are presupposed in our socially formed considered judgements about justice, with the result that in reflectively endorsing these principles we can restructure the basic structure of social cooperation of our social world. It is not just that 'justice must be seen in order to be

done' (Williams 1998: 246), but that justice must be seen, and reflectively endorsed, in order to *be* justice.

5.5 Constructivism as Realistic Utopianism

Cohen gets to the heart of his dispute with constructivism when he says that Rawlsians 'are in one way more Utopian than I', for constructivists 'believe it is possible to achieve justice', whereas for Cohen 'justice is an unachievable (although a nevertheless governing) ideal' (RJE: 254). Again, Cohen's utopian socialism is illuminated by comparing it to his previous realist historical materialism. Cohen and Rawls agree that it is Humean circumstances of justice – material scarcity and limited altruism – that give rise to the problem of justice and that, conversely, material abundance or unlimited benevolence would make justice unnecessary. If justice is a matter of compliance with institutional rules then a communist society in which material abundance makes possible the withering away of the state, will, Cohen recognises, be beyond justice. It will, in Rawls's terms, achieve justice 'without any reliance on people's sense of right and justice' (LHPP: 371, see also JF: 177). Now that the idea of such material abundance has been discredited two options remain. Cohen rejects the option, commonly though to be Marx's view, of socialized motivation, taking this to be akin to Hume's unlimited generosity in which a person feels no more concern for his own interest than for that of other persons. Instead, Cohen favours voluntarily just behaviour to produce a patterned egalitarian distribution required by justice.

In the widest sense of distribution there always is a distribution of goods and bads in society, and it always has a certain shape. Then where justice is a distribution of a certain shape, it is impossible to go 'beyond' justice and injustice (SFE: 143).

The reason Cohen now places justice beyond society is because he previously envisaged a society which thanks to material abundance was beyond justice.

Justice for Cohen is, then, an egalitarian distribution not a virtue of citizens (SFE: 141). Cohen takes Rawls to be concerned with rules of regulation not fundamental principles of justice. The former term is Cohen's rendering of Rawls's 'regulative principles', which Cohen disfavours because he takes it to misleadingly put one in mind of Kant's regulative ideas. Now Rawls indeed refers to our 'regulative sense of justice', a regulative sentiment that is regulative of our conduct – so far so Humean and un-Kantian (TJ: 498, 491, 514/436, 430, 451). But he later claims that on reflection this regulative sentiment will be found to be a 'regulative desire to act upon certain

principles of right' (TJ: 501/491), and his Kantian interpretation of justice as fairness compares the principles chosen in the original position with principles 'regulative of the kingdom of ends' (TJ: 256/226). Distinguishing between regulation and restriction (TJ: 203/178), Rawls can argue that principles of justice that are regulative of our sense of justice are constitutive of our autonomy and not a restriction upon it.

Rawls's constructivism and its synthesis of Humean and Kantian ideas is most extensively outlined in his *Lectures on the History of Moral Philosophy*, whose guiding thread is the problem of social theodicy. I will show that the lectures can be read as possessing a circular or dialectical structure, and this will help to explain the prefiguration of the final figure Rawls discusses in the lectures, Hegel, in the first, Hume. This is 'Rawls's Hegelian reading of Hume' (Lange 2009). Before examining this, we should remember the political concerns that lie behind Rawls's analysis of the 'methods of ethics'. The governing idea of a well-ordered society is one Rawls finds in Rousseau, the idea that we only fully realize our moral freedom in society, and, more specifically, in a society that realizes in its basic institutions a social compact that allows human beings to realize their potential to be guided by deliberative or public reason (LHPP: 231). This realizes a Kantian standpoint of self-legislation within Hegelian institutions. In Hegel, I will argue, the tension between Kantian ideal feasibility and Humean real feasibility is resolved in a realistically utopian conception of political philosophy.

Hume and Kant personify modern moral philosophy for Rawls. They are concerned not with a transcendent ethical good, but rather with a moral order that arises from human nature and the requirements of social life, and a world where the knowledge and motivation to act morally is accessible to every reasonable and conscientious person (LHMP: 11). Rawls begins with Hume, who offers 'an account of epistemology naturalized (to borrow Quine's well-known phrase) and of morality psychologised' (LHMP: 85). This sets the terms of Rawls's ensuing concern to defend a non-psychologised morality in a manner that is not inconsistent with a naturalized epistemology. Morality as a natural fact arises for Hume from the circumstances of justice. Individuals are not egoistic, but being 'realistic' we recognise that human generosity extends to only an immediate circle of family and friends. Rawls's own Hegelian social ontology is prefigured in his observation that Humean circumstances of justice might equally encompass the collective self-centredness of groups like universities and churches, which 'can be worse than individual egoism when supported by religious and philosophical doctrines' (LHMP: 59); reasonable pluralism is part of the circumstances of our social world. Hume pursues a practically best scheme of conventions for specifying property, 'given human beings as they are', one that is 'really

feasible' (LHMP: 64, 65). It is not 'the best scheme we can imagine, much less the best scheme allowing that human beings and our situation in nature might have been different' (LHMP: 60). This is Hume's fideism of nature, his aim of showing that 'morality is a natural phenomenon fully continuous with human psychology' (LHMP: 69).

Humean justice is a convention, regulated by a public rule, that expresses a common interest, and is supported by an idea of reciprocity. While self-interest is the original motive for the establishment of justice, the idea that compliance with rules is a virtue arises out of a sense of sympathy with the public interest, our sense of justice. Sympathy is transmuted into moral sentiment by our taking up the point of view of a judicious spectator. This testifies to our capacity for a counterfactual imagination: moral judgements 'may be founded not only on moral sentiments we *do* have (by actually taking up the judicious spectator's point of view), but *also* on what sentiments we *know* we *would* have *were* we to take up that point of view (LHMP: 89). A second Hegelian theme emerges in a prefiguration of the idea of reconciliation to our social world. According to Hume, 'not only virtue must be approv'd of, but also the sense of virtue: And not only the sense, but also the principles, from whence it is deriv'd' (Hume, quoted in LHMP: 100).

Hume is saying that his science of human nature also shows that our moral sense is reflectively stable: that is, that when we understand the basis of our moral sense . . . we confirm it is derived from a noble and generous source (LHMP: 100).

We could call this a hermeneutics of sympathy. It stands opposed to a hermeneutics of suspicion that undermines our common moral sentiments, Rawls mentioning the 'big three' of Marx, Nietzsche and Freud.

However, Rawls finds Hume's naturalistic account of the judicious spectator to be ultimately psychological and to lack a conception of practical reasoning. Hume fails to recognise that persons can act 'above the order of nature, in the sense that they can act independently of that order in the pursuit of personal and social ideals' (LHMP: 307). Principle-dependent desires can find a foothold in agents' subjective motivational sets, most importantly a desire to act according to reasonable principles of practical reason that regulate how a plurality of agents relate to one another. Having said this, Hume's conception of the judicious spectator can be developed into a conception of practical reasoning, a mutually recognised moral point of view that can be used as a criterion to work out, or construct, what our moral judgements should be. One route would be to

introduce Bentham's or Sidgwick's principle of utility as the fundamental principle of practical reasoning; Rawls, though, turns to Kant.

Rawls is therefore not so much proposing a Kantianism with a 'Humean face' (Sandel 1998: 14) as a Humeanism with a Kantian face. Kant for Rawls means to make us aware of, and arouse in us a desire to act on, a conception-dependent desire. A conception-dependent desire connects principle-dependent desires with a moral ideal. In Kant, this is a 'desire, belonging to us as reasonable persons, to act from an ideal expressible in terms of a conception of ourselves as autonomous in virtue of our free reason, both theoretical and practical' (LHMP: 148). Put positively, we have the capacity for a good will, and this capacity is a condition of our being members of a possible realm of ends, of being worthy of happiness. Rawls gives Kant's kingdom of ends a democratic Rousseauian interpretation: Kant's idea, 'basic to much democratic thought' was 'much influenced by Rousseau', to whose ideas Kant sought to 'give deeper philosophical expression' (LHMP: 160). Put negatively, though, 'unless we pursue our aims within the limits of the moral law, human life is worthless, without value' (LHMP: 158).

For Rawls, Kant's legacy for us as citizens of a constitutional democratic society is a constructivist conception of practical reason and a coherence account of its authentication. While the moral law applies to God and the angels as well as us, Rawls interprets the categorical imperative, which applies to reasonable beings, in independence from the moral law. Rawls's interpretation of the categorical imperative is at its most controversial in its idea of how the moral law must be brought 'nearer to intuition' in the Humean sense of being brought nearer to feeling. We need, says Rawls, a procedure by which the categorical imperative is applied to us as 'finite beings having a particular place in our social world and situated in the order of nature' in order to 'bring to life and to make intelligible Kant's characteristic deeper ideas' (LHMP: 164-65). This is the categorical imperative, or CI, procedure. Whereas Kant saw the three different formulations of the categorical imperative as equivalent, Rawls takes the formula of universal law, understood in a naturalized sense of a law nature, to be the most usable for us (LHMP: 183).

In following Dewey in overcoming Kantian dualisms through a naturalized Hegelianism, Rawls turns Kant's categorical imperative on its head by making the social rather than the individual primary. He does this first in seeing the categorical imperative as testing maxims for reconstructing the basic structure of our social world, not personal maxims (CP: 339). '[S]ince we are finite beings with needs, we cannot apply that imperative to our actions directly but can do so only after we have interpreted it in terms of the law of nature formula by setting out the CI-procedure' (LHMP: 182). The CI-

procedure generates an ‘as-if’ law of nature. Through our counterfactual imagination we adjoin this to existing laws of nature and consider how over time this would reform the order of nature, an ‘adjusted social world’ (LHMP: 169). In forcing us, in a Rousseauian sense, to view ourselves as legislators proposing public moral practice for a stable social world, the CI-procedure is analogous to the standpoint of our sense of justice represented by the original position, a conception of true human needs standing in for Rawls’s own social primary goods. But we must be able to will not only this adjusted social world but also to autonomously endorse it on the basis of our conception of the good. This question is pursued from the perspective of the third formulation of the categorical imperative, of autonomous legislators of a kingdom of ends. It is analogous to Rawls’s standpoint of deliberative rationality, which I discuss in the next chapter. Bringing the moral law nearer to feeling, so that we can form a conception of ourselves as members of such a world in a way that stirs our moral sensibility, requires the ideal of a realm of ends, ‘the conception of the social world that would come about (at least under reasonably favorable conditions) if everyone were to follow the totality of precepts that result from the correct application of the CI-procedure’ (LHMP: 224). Kant’s kingdom of ends fails to recognise the need for ‘the mutual recognition of the moral law in the public role of a society’s moral culture. Hegel will stress just this point’ (LHMP: 209). The second aspect of Rawls’s departure from Kant in emphasising the primacy of the social is therefore the requirement of full publicity central to Rawls’s constructivism (CP: 340).

Before turning to the demands of full publicity and then Hegel, however, Rawls contrasts the status of Kant’s moral law with Leibniz’s metaphysical perfectionism, understood as a species of rational intuitionism that stands opposed to Kantian constructivism.⁸⁶ In accordance with an orthodox Christian view, Leibniz presents ‘an ethics of creation’, specifying principles that ‘lie in God’s reason and guide God in selecting the best of all possible worlds’ (LHMP: 108). In *TJ*, Rawls says that on this conception ‘Even the general facts of nature are to be chosen’ (TJ: 159/137). Cohen glosses this as the idea that ‘the *function* of fact-independent principles is to determine what the general facts of nature are to be’, to which he responds: ‘they might have that function for God, but they need not therefore have that function for us’ (RJE: 261 n.50). But Rawls, unlike Cohen, is concerned with principles we can *will* as unconditionally good, and Cohen in no way contests Leibniz’s view that the unconditionally good is a *perfect* world, requiring the specification of principles from a God’s eye point of view (even if Leibniz thinks we know the principles of God’s reason whereas Cohen thinks

⁸⁶ Here I follow the order in which the lectures were delivered rather than the order in which they are published. See LHMP: 105.

knowledge of fundamental principles is unachievable). We may note in passing that Leibniz's neo-Platonic predicate-in-subject theory of truth has a pivotal place in the history of the conception of analysis practiced by Moorean analytic philosophy (Beaney 2009). If one were to speculate about the transcendent Platonism of much analytic normative philosophy, one could start from Rawls's observation that 'Leibniz's theory of truth is framed for his philosophical theology and its apologetic aim' (LHMP: 122).

Rational intuitionism makes basic moral concepts unanalyzable as they are conceptually independent of natural concepts, outside the order of nature. Rawls argues that Kant's conception of autonomy is such that Kant would reject rational intuitionism as a form of heteronomy just as much as Hume's psychological naturalism. Any proposed independent object – any 'exalted entity' (SGEK), whether Platonic, Leibnizian, or naturalistic – must 'first stand trial . . . at the bar of the supremely authoritative principles of pure practical reason before we can grant its reasonableness' (LHMP: 229). Rawls therefore prefers the strategy of post-Kantian Hegelians to Kant's own strategy for explaining how pure reason could be practical through a 'fact of reason'. Rawls argues that the constitution of reason must be seen as self-authenticating: 'reason cannot be judged or explained by anything else' (LHMP: 271). This is the Quinean-Hegelian holism of Rawls's reflective equilibrium. Rawls suggests that Kant himself gave up the fact of reason in favour of a coherence account of the authentication of practical reason, but this comes with a warning label attached: 'you should know that this interpretation of [Kant's] thought is not generally shared. You must regard it with caution' (LHMP: 268).

Practical reason, then, 'constructs for the will its own object out of itself', the public moral order of a possible realm of ends, 'and does not rely on a prior and antecedent order of values' (LHMP: 226). It must be reasonable to try to realize the ideal of a possible realm of ends in the natural world, at least under reasonably favourable circumstances, and the ideal of a realm of ends is not a social world that can be described prior to and independently of the concepts and principles of pure practical reason and the procedure by which they are applied. Rawls's account of moral objectivity is ultimately Rousseauian in identifying objective moral convictions as those backed up with 'reasons sufficient to convince all reasonable persons' (Neuhouser 2009: 213 n.40; LHMP: 245). Real reasons are the 'public reasons – reasons people freely present in good faith to one another' (LHMP: 302). Here is where Kant's religious view of human nature lends his thought, in Rawls's view, a Rousseauian cast. It is part of human nature to be able to act above nature in a moral community of mutuality and self-respect, an aristocracy of all.

We end, finally, with Hegel. Following post-Kantian Hegelians, Rawls, in his second lecture on Hegel, rejects Hegel's conception of *Geist* as self-awareness of the highest good that transcends the knowledge of individuals. More interesting, though, are the Hegelian ideas that Rawls accepts in his first lecture, which addresses 'how Hegel thought the concept of freedom was actually realized *in* the social world through political and social institutions at a particular historical moment' (LHMP: 330). Rawls interprets Kant as arguing that the moral law determines elements of the *matter* as well as the form of pure will (LHMP: 194). But Kant neglects the point that Rawls takes to be crucial that the 'matter' must include a scheme of social institutions. The free will must have itself as its object, must will itself as the free will:

The free will wills itself as the free will, first, when it wills a system of political and social institutions within which it can be free . . . when, second, in willing the ends of those institutions it makes their ends its own, and, third, when it is thereby willing a system of institutions within which it is *educated* to the concept of itself as a free will by various public features of the arrangement of those institutions, features which exhibit the concept of a free will (or of freedom). Note here the significance of education (*Bildung*) (LHMP: 338).

If we can affirm the social world as constitutive of our freedom – if it can appear justified to free thinking – we can become reconciled to our social world. The potential realizability of reconciliation confirms that the good of a realm of ends is 'one that can be reasonably striven for (although never fully realized)'. In doing so 'we are not being visionaries who lack a sense of realism' (LHMP: 225). However, in order for the modern social world to be a 'home', for us to be able to be at home in it, it must be possible for people to actualize themselves both as individuals and as social members, reconciling the liberties of the moderns and the liberties of the ancients (Neuhouser 2000: 129). We will see how Rawls seeks to address this problem in Part III.

Conclusion

I have argued in Part II that Rawls's project of reconstructing the basic structure of our social world involves an internal relation between the principles of justice chosen in the original position and the general facts of our social world. The role of a conception of justice is to systematise our considered judgements and a conception of justice is ultimately justified by its capacity to reform the basic structure of our social system of cooperation in a manner acceptable at the tribunal of the reason of free and equal citizens. Cohen argues that individuals must account for themselves at the 'bar of

justice' (Cohen 2008b: 13), but Rawls follows G.E.M. Anscombe (1958b) in taking such judgement – whether by God, or the concept of justice – to be an anachronistic law conception of ethics that should be anathema to modern moral philosophy. For both Rawls (CP: 286-303) and Anscombe, progress in moral philosophy depends upon progress in moral psychology.

Cohen's critique of Rawls purports to be an internal critique based on a fundamental meta-ethical truth about the relationship of facts and principles. However, on closer inspection it turns out to depend upon attributing to Rawlsian constructivists an almost sectarian clarity of mind and analysing this belief structure in accordance with a philosophical methodology that Rawls rejects. If on reflection constructivists find they lack such clarity of mind they may agree with Rawls that the lack of transparency of our reason may lead us to think that there are moral principles 'prior to and independent of reason', but that we should dispel this illusion and recognise that 'we are really using our reason to describe itself' (PL: 243). Requiring of principles of justice that they be public and stable is after all Rawls's attempt to take to heart Marx's warning regarding the descent of normative thinking into ideology.

Moore took it that we may always ask whether what is described as good really is good. If someone were to complain that philosophy ought rather to be practical and constructive, Moore could only reply: 'If any one does not care for knowledge for its own sake, then I have nothing to say to him' (Moore 1993 [1903]: 115). In like manner, after 'rescuing' the concept of justice, all Cohen (2009b) can say is that 'justice just is justice'. In the same way that one can plausibly see Moore's intuitionism as sowing the seeds of emotivism (MacIntyre 2007), so Cohen's positing of principles of justice that transcend the facts of this world places justice beyond reasonable belief, lacking any basis for sustaining our reasonable faith in common human reason. Analytic philosophy threw out the baby of internal relations with the bathwater of Hegelian metaphysics, leading Cohen to dismiss the Hegelianism of Marx and ignore the Hegelianism of Rawls. In neither case does Cohen recognise the need to 'constrain theory to play an emancipatory role for the very social world that is its object, through that world's acceptance of the self-interpretations the theory offers it' (Sensat 2003: 323). The resources of the rational agency tradition are not exhausted, and consist in the search for an

orientation to the social world and its historical situation that avoids two abdications of responsibility: utopianism, whether pious or fanatical, on the one hand, and fatalism, whether complacent or cynical, on the other. Philosophy's task is to vindicate or defend this possibility, to provide reassurance of our competence (Sensat 2007: 1-2).

PART III. RAWLS'S REALISTIC UTOPIANISM

6. The Problem of Self-grounding in the Search for Stability

In Part III, I will analyse and critique the development of Rawls's constructivist project from the perspective of its aspiration to realistic utopianism. Simplifying somewhat, we can divide Rawls's project into the early Rawls of *TJ* and the late Rawls of *PL*, with Kantian Constructivism an intermediate stage (Freeman 2003a). Having considered Rawls's constructivism in broad terms, my concern now is to set out the fundamental continuity of Rawls's project and the light this sheds on Rawls's 'political turn'. The publication of Rawls's complete *oeuvre*, including his undergraduate thesis, has made it possible for the first time to assess Rawls's project as a whole. As David Reidy puts it, with some asperity, 'not every debate over Rawls's political philosophy has been well motivated by Rawls's arguments' (Reidy 2008).⁸⁷ My starting point is Anthony Laden's argument against the 'standard blueprint' for reading Rawls and in favour of an 'alternative blueprint'. In seeking to uncover the roots of Rawls's political liberalism in Rousseau and Hegel, Laden has drawn attention to the limitations of pure Kantian interpretations of Rawls. I have already discussed a number of theorists who interpret Rawls from the perspective of the Kantian paradigm and a number of others are discussed in this chapter. The alternative blueprint threatens, however, to replace one partial perspective with another. So far we lack a view of Rawls as a complete thinker.

In seeking such a view, it is illuminating to consider the reasons Rawls gives for interpreting J.S. Mill as having broken with a liberalism of the good and adopting a liberalism of freedom. This lies in Mill's recognition of the need for a Coleridgean counterweight to compensate for Bentham's defects as a 'half-thinker'. Political principles for the modern democratic age 'must not be concerned primarily with how to provide legal incentives for good conduct', but rather with arranging basic institutions so that citizens come to have a character that is conducive to good conduct (LHPP: 255). Mill endorses a division of justificatory labour in which 'moral doctrines contain normative concepts and principles combined with elements of human psychology and political sociology, together with other institutional and historical assumptions' (LHPP: 270, 313). Coleridge said one was born either a Platonist or an Aristotelian, but for Mill the problem of the modern age was the division between Benthamites and Coleridgeans (Mill 1987 [1840]: 180). Rawls follows Mill in seeking to combine the complementary strengths of the methods of both: 'To Bentham it was given to discern more particularly

⁸⁷ Reidy (2009) is insightful about the role of biography in structuring but not determining philosophical interpretation.

those truths with which existing doctrines and institutions were at variance; to Coleridge, the neglected truths which lay *in them*' (Mill 1987 [1839]: 133).

I will argue that situating Rawls within the critical theory framework of the problem of self-grounding offers the best prospect of a picture of Rawls as a complete thinker. Rawls argued for the independence of moral theory (CP: 286-303), which integrates moral conceptions with psychology and social theory, from the theory of meaning, epistemology, the philosophy of mind, and metaphysics. This is strikingly similar to Max Horkheimer's (1972 [1937]) rejection of traditional theory in favour of a critical theory that is independent of first philosophy and science. Although commonly presented as quintessentially an analytic philosopher (see e.g. Pettit 1993), '[b]oth methodologically and stylistically [Rawls] departs from the analytic tradition' (Freeman 2007a: 28).⁸⁸

I first set out a third self-grounding constructivist blueprint for reading Rawls. Then, beginning with Rawls's early work, I consider the way in which the elements of justice as fairness are concatenated, and therefore how they hang together, in Sellars's broadest possible sense of the term, in *TJ*. Next, I consider these elements as part of a developmental system, interpreting Rawls's attempt to resolve the problem of stability in part three of *TJ* within the framework of the problem of self-grounding and outlining how this leads to *PL*. We will then be in a position to turn in the next chapter to an assessment of Habermas's critique of Rawls.

6.1 A Self-grounding Constructivist Blueprint for Reading Rawls

In assessing 'The House that Jack Built', Laden (2003) argues that the interpretative community that has grown up around Rawls's construction is a house divided (see also Matan 2004). According to the 'standard blueprint', Rawls is engaged in a grand philosophical project, developing a 'theory in the traditional sense of the word', and, particularly through the argument from the original position, concerned to demonstrate the rationality of justice (Laden 2003: 371). By contrast, according to the 'alternative blueprint', Rawls is concerned with a narrower project of providing public justifications through public reason, pursuing a conception of 'philosophy as defense' rather than traditional philosophical theorizing (2003: 379). Rawls's philosophical defence of the possibility of a just constitutional democratic regime contrasts with the '*modus operandi*

⁸⁸ *TJ* 'simply bypasses . . . metaethical issues . . . It is a book which descends straight from Kant, Mill, and Sidgwick. The same book could have been written if logical positivism had never existed' (Rorty 1982), although not perhaps if the Quinean postanalytic reaction against logical positivism had not occurred. See Alexander (1974). Rajchman (1985) presents Rawls as a postanalytic philosopher.

of a theoretical approach [which] is to develop a theory and then apply it' (Laden 2004: 288). Laden subscribes to the alternative blueprint, and follows Rousseau in understanding the philosopher as a citizen addressing fellow citizens (2004: 289 n.31), the authority of philosophy being 'simply the authority of democratic politics' (Laden 2001: 5-6).

I will conclude Part III by arguing that Laden's alternative blueprint has its own limitations, but it plays a valuable initial role in drawing attention to the inadequacies of pure Kantian interpretations of Rawls. It might seem strange, then, that central to Laden's own view is an appeal to Korsgaard's conception of practical identity, albeit reinterpreted in the terms of the politics of identity. But this is suggestive of how far Korsgaard's Fichtean Kantianism takes a step in a post-Kantian direction. The problem for a third self-grounding blueprint for reading Rawls, I will argue, is to develop a view of constructivism that combines Korsgaard's concern with autonomy and self-grounding with O'Neill's concern with political objectivity.⁸⁹

Korsgaard provides a helpful characterisation of Rawls's constructivism, the *concept* of justice as a problem of social cooperation between free and equal persons setting the parameters to which the *conception* of justice as fairness offers a practical solution by reflecting on the nature of the problem itself (Korsgaard 2008: 322). But Korsgaard's own approach offers limited resources for developing a critical theory. Whereas Korsgaard focuses on the normative implications of the practical identity of the self, Rawls describes a 'practical identity' between expressing our nature as free moral persons and the desire to act justly according to the principles of justice chosen in the original position (TJ: 572/501). Korsgaard collapses back into the identity of the individual moral agent capacities that Rawls argues social institutions must develop and stabilise. Whereas Korsgaard understands constructivism as a method to be developed separately in moral and political philosophy, I have taken the problem of self-grounding to be *the* modern problem of autonomy such that we cannot take these disciplinary demarcations for granted.

Korsgaard's Fichtean Kantianism can claim to be a 'deep' constructivism, one which goes 'all the way down' (Korsgaard 2008: 324; Watkins 2002: 358). Even our most basic reasons can be constructed, since as agents we determine for ourselves whether a consideration in favour of doing something should be treated as a reason. But this comes at the cost of endangering the objectivity to which a political constructivism might justifiably aspire. In emphasising objectivity rather than autonomy, the problem

⁸⁹ Much of the literature on constructivism conceives of it in broadly defined terms and therefore does not always illuminate Rawls's view. See Brink (1987); Hill (1989); Klosko (1997); Roberts (2007); McKinnon (2002).

for O'Neill's Kantian constructivism, on the other hand, is whether it *can* go 'all the way down'. O'Neill (2003a) thinks it should so aspire because she is critical of the fact that Rawls, despite his constructive account of the content of practical reason in the original position, offers only a 'coherence' justification of the original position itself. The justification of practical reason is left to rest on unvindicated considered judgements rather than universal principles. Although I cannot pursue it in detail here, there are grounds for questioning whether O'Neill's (1992) strategy for overcoming the foundationalist problem of Münchhausen's trilemma by a constructivist vindication of practical reason is successful. The coordination of a plurality of rational ends is given by her only an instrumental or teleological justification, serving as a goal or end rather than a rationally inescapable universal presupposition (Budde 2009). O'Neill (1990: 218) takes Rawls's constructivism to teeter on a knife edge between idealization and relativism, but I find no grounds for revising my conclusion that it is pure Kantianism that is unstable and that the real choice is between traditional metaphysical Kantianism and post-Kantianism.⁹⁰

O'Neill's interpretation of Rawls's constructivism suffers from its acceptance of the basic contention of Sandel's (1996) communitarian argument that a Kantian conception of the person provides the 'foundation' of Rawls's theory. O'Neill (1988, 1996, 2006) argues that the original position includes an idealized as opposed to abstract conception of practical agency, one that in its model of persons adds false premises rather than bracketing inessential ones.⁹¹ However, the moral powers of citizens represented in the original position – the capacities for a reasonable sense of justice and a rational conception of the good – may be compared to Hegel's description of modern free personality in terms of the capacities of 'self-conscious universality' and 'self-determination' (Schwarzenbach 1991: 551). Rawls, like Hegel, abstracts from various *existing* features of the modern social world, but would deny that this is abstraction from reality (Hardimon 1994: 58). The original position idealizes insofar as in identifying the basis of persons' capacity for autonomy it uncovers the potential for reconceiving the structure of the relationships between persons within and towards their social world. But this idealization is justified insofar as it informs the choice of principles to reform the basic structure of the social world in a manner that may be reflectively endorsed by free and equal citizens.

⁹⁰ Forst (2002) remedies a number of the deficiencies of O'Neill's view by integrating it with Habermas's discourse theory. I cannot do justice to this view here, but it does not succeed in resolving all ambiguities about how contexts of justice defined by the universality of scope of practical reason relate to contexts of justice circumscribed by the social structures of modern democratic societies.

⁹¹ See also Schwartzman (2006); Mills (2005); Pateman and Mills (2007). For discussion, see McCarthy (2004); King (2008).

I will present the problem of self-grounding that Rawls's constructivism faces as a matter of achieving reflective equilibrium between the normative grounds of the original position and their grounding in their capacity for reflective endorsement. Rawls's constructivism seeks to develop its justification out of itself, to find in its grounding the justification of its grounds. A conception of justice constructed in such a manner can aspire to form citizens who can reflexively recognise the justification of the constructivist procedure through which that conception was constructed. The grounds of practical reason are reflexively justified in finding an immanent grounding in the reflective endorsement by citizens of the conception of justice that practical reason has constructed within their conceptions of the good.

In analysing Rawls in this way, we can draw upon James Bohman's (2005) distinction between feasibility and realizability. Rawls and Habermas 'both accept that democratic ideals are constrained by the social facts of "pluralism" and "unavoidable social complexity"' (Bohman 1996: 10). But a critical theory asks not only what facts are salient to theory construction, but also how these facts are to be normatively conceived. Feasibility involves taking into account facts as such: they appear simply as constraints. Realizability, by contrast, takes into account how such facts should be understood in relation to human agency: how they have been entrenched by historical processes, in what respect they are subject to change, how facts exist as facts. Laden recognises that his own deliberative liberalism 'overlaps in many ways' with Bohman's more Habermasian pragmatic critical theory (Laden 2001: 107 n.15), and Bohman recognises that Rawls can be seen as a critical theorist in a broad sense (see also Kelly 2001; Ron 2006). Attuned to the Rousseauian and Hegelian dimensions of Rawls's political liberalism, as well as Rawls's realizability conception of the relationship of democratic ideals to general facts of our social world, we can now read Rawls within the framework of a self-grounding constructivist blueprint.

6.2 The Elements of Justice as Fairness

In assembling the elements of justice as fairness, I begin with reflective equilibrium, which provides the framework within which Rawls's other conceptions of justification are situated. Reflective equilibrium appears in embryonic form in Rawls's first article, 'Outline of a Decision Procedure for Ethics' (CP: 1-20). Rawls argues that the objectivity of moral values depends not upon independently existing ideal entities, but upon a reasonable constructive ethical procedure in which competent judges adjudicate between competing interests. Such a method is analogous to the understanding of

objectivity in science, making ethics similar to the study of inductive logic. Principles are sought to account for experience, and are then tested by their capacity to accommodate future experiences. Rawls's procedure is circular in the sense that there are no non-moral criteria by which to select the competent judges who are to select moral principles, but not therefore 'valueless' as was immediately claimed (Mardiros 1952: 225). Its account of the mutual adjustment of principles and considered judgements parallels Goodman's account of the 'constructive' mutual adjustment of principles of deductive and inductive inference (TJ: 20 n.7/18 n.7). As Goodman says, 'this circle is a virtuous one': '*A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend*' (Goodman 1979: 64).

Rawls's early development of reflective equilibrium has, however, been taken to contradict the idea that reflective equilibrium is based on the holistic epistemology of Goodman, Quine and White (Freeman 2007a: 64). But this is misleading on two fronts. First, it distracts attention from the epistemological ideas upon which Rawls *does* draw in this early article, his conception of explication, or rational reconstruction, putting one in mind of Rudolf Carnap, upon whom Rawls draws quite extensively in the doctoral thesis from which the article derives (SGEK). While working within the tradition of analytic philosophy, from the start Rawls understands 'analysis' as 'transformative' rather than 'decompositional', as replacing existing conceptions with new conceptions that exceed them in clarity and systematicity, rather than analysing existing conceptions in an effort to isolate their intrinsic nature (Beaney 2007).

Second, Rawls develops reflective equilibrium in a holistic and pragmatist direction.⁹² In *TJ* there is an erosion of the positivist distinction between observation and theory that testifies to the influence of Quine and contributes to a more ambitious view of the potential revisability of common sense considered judgements (Delaney 1977; see also Daniels 1996: 43 n.16; TJ: 579 n.33/507 n.34). The historically shaped and socially stabilised considered judgements that postfoundationalist reflective equilibrium takes as

⁹² Although note that in *TJ* Rawls says that he is returning to the conception of moral theory set out in 'Outline' but not adopted in the intervening essays (TJ: xi/xx-xxi). Rawls accepts Care's (1969) criticism that these essays took a contractarian procedure to defend a theorem – a necessary truth – about what principles persons will acknowledge as deductively following from the conditions in which they are situated, thereby making the nature of morality logically prior to a contractarian procedure. Care argues, and Rawls accepts, that the acceptance of principles should instead be seen as following non-deductively from deliberations that characterise ordinary practice, an activity of substantive moral criticism rather than logical analysis. Rawls credits Dreben for helping him to avoid this objection, persuading him of the Quinean view that 'notions of meaning and analyticity play no essential role in moral theory'.

its object should not be confused with ‘intuitions’.⁹³ When Rawls says ‘There is a definite if limited class of facts against which conjectured principles can be checked, namely our considered judgements in reflective equilibrium’ (TJ: 44/51), he means that the small but definite class ‘emerges only when reflective equilibrium is reached, and is still revisable in the light of further theory change’ (Daniels 1996: 43 n.17). Rawls is in the same boat as Quine; that is, Neurath’s boat, to which Quine (2001 [1969]) refers in observing the circular or reciprocal containment of epistemology in natural science and natural science in epistemology.

Rawls develops this method of ethics in an attempt to surpass utilitarianism. But his attitude to classical utilitarianism is ambivalent and he does not simply dismiss it as an ‘incorrect theory’.

Utilitarianism historically goes together with a coherent view of society, and is not simply an ethical theory, much less an attempt at philosophical analysis in the modern sense. The utilitarian principle was quite naturally thought of, and used, as a criterion for judging social institutions (practices) and as a basis for urging reforms (CP: 33 n.21).

If we aspire to systematicity, philosophical counter-examples to utilitarian reasoning should not simply lead to ad hoc intuitionistic trade-offs between utility and other values. Before turning to the social contract tradition to develop a systematic alternative, Rawls sought to show how a form of rule utilitarianism could overcome traditional objections to utilitarianism. In doing so, he developed the view that the primary subject of social justice, within the context of which reflective equilibrium will operate, is the rules of social practices. In ‘Two Concepts of Rules’ (CP: 20-47), Rawls distinguishes the justification of a practice or an institution from the justification of a particular action falling under a practice or institution.⁹⁴ This requires distinguishing a practice conception of rules from a summary conception. On the summary conception, ‘rules are pictured as summaries of past decisions arrived at by the *direct* application of the utilitarian principle to particular cases’ (CP: 34). On the practice conception, by contrast, rules define a practice and it is the rules themselves that are the focus of utilitarian principles. The rules of a practice are ‘logically prior to particular cases’ (CP: 36), constituting the actions persons can intelligibly perform within practices.

⁹³ The failure of critics such as Hare (1989) and Singer (1974) to acknowledge this can be traced to their ‘latent positivism’ (Daniels 1996: 43 n.17). ‘This whole methodological dispute has much more to do with one’s views on the analytic-synthetic distinction . . . than Hare is willing to admit’ (Delaney 1975: 105). However, for the view that reflective equilibrium is consistent with foundationalism see DePaul (1986), and that it is a modest foundationalism, Ebertz (1993).

⁹⁴ Rawls employs the term practice as ‘a sort of technical term meaning any form of activity specified by a system of rules which defines offices, roles, moves, penalties, defenses, and so on, and which gives the activity its structure’ (CP: 20 n.1).

This suggests a solution to two problems, one concerning the justification of a practice and another concerning the justification of an action within a practice. Regarding the former, when it comes to the institution of punishment, Rawls argues that the pursuit of utility-maximization should be restricted to legislators, who adopt a future-oriented perspective, while judges, who work within the system of rules, should abjure utilitarian reasoning. Regarding the latter, when it comes to the institution of promising, a promisor must take into account not simply the value of keeping an individual promise, but also the utilitarian social value of having a practice of promise-keeping. The point of the practice of promise-keeping, Rawls argues, is to renounce one's claim to make utilitarian calculations in the overall interest of successfully coordinating plans. Rawls's focus on the rules of social practices is important because it provides insights into his view of philosophy as 'constructing justice for existing practices' (James 2005). This understanding of social institutions is, as we saw in the last chapter, based on Hume's account of justice, but is joined with Wittgenstein's idea of the rules of a practice (Freeman 2007a: 23).⁹⁵ There may be a sense in which this is a conservative or quietist conception of philosophy (Cladis 1994; Mondal 2008), but it is not conventionalism: 'One can be as radical as one likes but in the case of actions specified by practices the objects of one's radicalism must be the social practices and people's acceptance of them' (CP: 43).

When Rawls begins to develop his substantive conception of justice as fairness in the article of that name (CP: 47-73) he ask his readers to 'Imagine a society of persons among whom a certain system of practices is *already* well established' (CP: 52). Rawls associates with Kant the principle of equal liberty he first articulates in this article, the idea that 'each person participating in a practice, or affected by it, has an equal right to the most extensive liberty compatible with a like liberty for all' (CP: 48). However, a more proximate influence is H.L.A. Hart. A principle of equal liberty is a *presupposition*, in Hart's (1955) sense, of ascribing to persons any particular rights. An institution is just when it satisfies a principle of equal liberty because this is a principle which participants 'could propose to one another for mutual acceptance in an original position of equal liberty' (CP: 77). Moves in the 'game' of a social practice are fair if they accord with rules that counterfactually could have been agreed to in a fair initial situation.

Far from being a 'refined Hobbesianism', Rawls's method seeks to derive Kantian conclusions from a Humean starting point. It is only on the basis of the former mistaken

⁹⁵ This Humean view is mediated by H.L.A. Hart's (1997 [1961]) account of law in terms of conventions (Freeman 2007a: 103). Reidy (2009) observes that 'Rawls thought of Wittgenstein as the most significant influence on his life's work'.

supposition that Barry could criticise the way *TJ* purportedly gave a new slant to ‘Justice as Fairness’, moving from a view of ‘morality as a system of mutual self-defence’ to one where the desire to act justly is ‘a central aspect of human development, a natural (and reflectively supported) extension of love for particular people and loyalty to particular associations’ (Barry 1973: 2; see also Buchanan 1972). However, the fact that the second chapter of Part three of *TJ* is based on ‘The Sense of Justice’ (CP: 96-117), published five years after ‘Justice as Fairness’, suggests a process of architectonic construction not visionary change. The sentimental moral psychology of a ‘sense of justice’ can seem philosophically suspect alongside the instrumental rationality of social choice theory. But quite how fundamental it is to Rawls’s project is evident from Rawls’s remark that ‘[i]n the absence of a sense of justice on everyone’s part, there would be . . . no objection to the utilitarianism principle’ (CP: 115). When we take persons to have a sense of justice we need no longer appeal to a Hobbesian sovereign, nor to a utilitarian technocrat, to conceive of the stability of a practice of social cooperation (CP: 104), but can appeal instead to the moral psychological mechanisms examined by Rousseau and Hegel (CP: 426 n.10).

Already present in Rawls’s early rule utilitarian proposal is the idea that utilitarianism is self-effacing when persons engaged in a practice of cooperation act directly on a principle of utility maximization. A practice of promise-keeping could not exist if persons maintained a utilitarian proviso to break promises. Similarly, an institution of punishment would not exist – would not fulfil our criteria for punishment – if it were public knowledge that the authorities ‘punished’ suspects on utilitarian grounds. The self-effacing tendency of utilitarianism led Sidgwick to suggest that utilitarianism would have to remain an esoteric morality, a system from which the ‘vulgar should keep aloof’ (Sidgwick 1981 [1907]: 489-90), or what Williams later criticised as ‘Government House utilitarianism’ (Williams 1993: 108-110). Rawls seeks principles of justice which, like utilitarianism, take as their end the public good, but unlike utilitarianism achieve stability on the basis of a publicly available justification, principles generating their own support. The problem with utilitarianism is its conception of impartiality, represented through the idea of a rational and impartial sympathetic spectator. When the approvals of the impartial sympathetic spectator are adopted as the standard of justice, ‘this results in impersonality, in the conflation of all desires into one system of desire’ (TJ: 188/164). This suggests an administrative conception of justice in which an ideal legislator seeks to maximize the overall sum of net satisfaction in the same way as an entrepreneur seeks to maximize profit (TJ: 27/24). The standpoint of the impartial spectator is not one that can be adopted by citizens because its moral ideal is one of perfect altruism, and ‘a perfect altruist can fulfil his

desire only if someone else has independent, or first-order, desires' (TJ: 189/165). If everyone were a perfect altruist no one could be an altruist; a problem of justice requires at least two persons with conflicting interests. There is therefore an inevitable and ineluctable disjunction between utilitarianism's description of individuals' preferences and the prescriptive theory according to which those preferences are organised into a system.

Rawls famously concludes in *TJ* that 'Utilitarianism does not take seriously the distinction between persons' because it adopts 'for society as a whole the principle of rational choice for one man' (TJ: 27/24). However, at first sight this seems to flatly contradict Rawls's earlier observation that the 'assimilation of justice to a higher order executive decision' brings out utilitarianism's 'profound individualism, in one sense of this ambiguous word' (CP: 66). The apparent contradiction can be resolved by distinguishing between ontological separateness (individualism) and moral separateness (the moral distinctness of persons). In treating persons as ontologically separate sites of pleasure and pain to be ordered by an ideal legislator, utilitarianism ignores the moral distinctness of persons, the Kantian idea of individuals as morally autonomous, according to which the problem of justice can be reconceived as one of collective self-legislation. Rawls nowhere expresses his understanding of the moral distinctness of persons more clearly than in his undergraduate thesis: 'Salvation *integrates personality into community*, it does not destroy personality to dissolve it into some mysterious and meaningless "One"' (MSF: 126). In the terms of *TJ*

a love of mankind that wishes to preserve the distinction of persons, to recognize the separateness of life and experience, will use the two principles of justice to determine its aim when the many goods it cherishes are in opposition. This is simply to say that this love is guided by what individuals themselves would consent to in a fair initial situation which gives them equal representation as moral persons (TJ: 191/167).⁹⁶

Utilitarianism is not *wrong*, but rather out of step with the times, or at least will be shown to be if justice as fairness succeeds. The argument that when represented as free and equal persons we would not choose one of the historically dominant conceptions of justice, like the principle of average utility, but rather the principles of justice as fairness, is an argument that utilitarianism is anachronistic because justice as fairness better systematises our considered judgements. Whereas justice as fairness 'can account for the utilitarian insight, the reverse is not the case' (Schwarzenbach 1991: 567). This results in a reconfigured economy of philosophical ambition as a result of placing autonomy not

⁹⁶ On the role of love in Rawlsian stability, see Mendus (1999).

utility at the heart of a method of ethics. The parties in the original position represent our natural right to equal liberty, in Hart's sense, in the procedure for choosing principles for reforming the basic structure of our social world. But if justice as fairness is more modest than utilitarianism in not offering a blueprint for reforming technocrats, it is more ambitious in its aspirations to reform citizens' self-understanding. While with utilitarianism 'the place we start from may always influence the path we are to follow', justice as fairness is not at the mercy of existing wants and interests (TJ: 263/231). Rawls is not constructing justice for existing practices, but seeking to reconstruct our practices in accordance with a conception of justice.

6.3 A Theory of Justice *In Mediis Rebus*

Reflective equilibrium, the basic structure as a social practice of social practices, the liberty principle, the original position and the sense of justice all find a place in *TJ*, a work of great architectonic complexity in which form cannot be separated from content in establishing meaning. Reading *TJ* in accordance with a self-grounding blueprint means, as urged by Dreben, beginning *in mediis rebus*: 'our moral theorizing begins in the messy middle of things, pursuing a holistic justificatory strategy of assessing principles in terms of the principles that we already hold' (Ridge 2001: 476).⁹⁷ It also means recognising the circular, indeed dialectical, structure of *TJ*, Part three justifying the starting point of Part one rather than following as a deductive conclusion from first principles.⁹⁸ Rawls warns that 'without consideration of the argument of the last part the theory of justice will be misunderstood' (TJ: ix/xviii).⁹⁹ 'The argument for the principles of justice is not complete until the principles selected [in the original position] are shown . . . to be sufficiently stable; and doing this extends to the next to last section' (PL: 140-41 n. 7). The three parts 'are intended to make a unified whole by supporting one another' (TJ: 579/507). Justice as fairness 'cannot succeed' if from a *seemingly* individualistic starting point it cannot explain the value of community (TJ: 264/234).

A theory of justice is tested by 'how well it introduces order and system into our considered judgments' (TJ: ix/xix). Rawls here announces his intention to proceed immanently, *introducing* system into our considered judgements through the use of simplifying devices whose conditions we do, or can on reflection, accept. It is a question

⁹⁷ See more generally Baker (2001). In recognising the extent to which Rawls shares Dreben's conception of philosophy – see Rawls (2001b) – one must remember that Rawls's conception of democracy drives his conception of philosophy and not vice versa (Laden 2004: 289 n.31).

⁹⁸ Schuldenfrei (1972) makes a similar argument regarding Quine's *Word and Object*.

⁹⁹ This is the part Rawls 'liked best' (Rawls 1991) and thought one of his most original contributions (Freeman 2003b: 143 n.2).

of how, as Rawls would later put it, the idea of society as a fair system of social cooperation can be ‘*unfolded* so as to find principles specifying the basic rights and liberties and the forms of equality most appropriate to those cooperating, once they are regarded as citizens, as free and equal persons’ (PL: 27, emphasis added).

Freeman describes the three parts of *TJ* as responding to Kant’s problem of how a just constitution is possible, requiring ‘first, the “correct concept” of a just constitution; second, “great experience during much of the world’s course”; and third, “above all else a good will prepared to accept that constitution”’ (Freeman 2007b: 145). There is something to this, but Rawls’s appeal to Kant again needs to be set in the full context of the RKH line. We could more properly compare Rawls’s task with that of Rousseau’s Legislator, first to understand principles of political right, second to discover institutions which embody these principles, and third to gain the consent of people to follow these institutions (Kelly 1987: 322). ‘As with *The Social Contract* and *Emile*, Rawls’s statement of his theory proceeds in two stages’ (Chapman 1975: 588). These two stages are, however, mediated by a conception of social institutions. We can think of the problem of *TJ* as that of realizing a Rousseauian general will within Hegelian social institutions in accordance with the conception of the autonomous person that, for Rawls, Kant showed to be the foundation of the general will. Rawls sets out four questions that must be considered by ‘any political conception of right and justice, including Rousseau’s’ (LHPP: 237). As we will see, the first three correspond to the three sections of *TJ*, the third having two sub-questions, while the failure of *TJ* to adequately address the fourth leads to *PL*:

- (1) What are reasonable principles of justice, and how is their correctness established?
- (2) What political and social institutions realize these principles?
- (3) How do people (i) learn principles of justice and (ii) acquire the motivation to act from them so as to preserve stability over time?
- (4) How might a society realizing principles of justice come about?

Methodologically speaking, the essential elements of Part one of *TJ*, ‘Theory’, are the introduction of the original position in relation to the role of principles of justice in regulating the basic structure of the social system and the justification of the conditions modelled in the original position through reflective equilibrium. The veil of ignorance allows the original position to play a role analogous to the social pact in Rousseau’s *Social Contract* (Dent 2005: 228). For Rousseau, in ‘the total alienation of each associate with all of his rights to the whole community . . . since the condition is equal

for all, no one has any interest in making it burdensome to the rest' (Rousseau 1997 [1762]: 50). In hypothetically placing ourselves in the original position we are 'forced' to think in terms of the general will, or public reason. The formal constraints of the concept of right – the parties must choose 'a set of principles, general in form and universal in application, that is to be publicly recognized as a final court of appeal for ordering the conflicting claims of moral persons' (TJ: 135/117) – follow from the problem of self-legislation to which the social pact is the solution.

Part two, 'Institutions', shows first how the standpoint of justice for the ordering of the basic structure modelled by the original position can be realized at different levels within the institutions of a constitutional democracy, the implications of justice as fairness being tested by its capacity to systematise our considered judgements at all levels of generality. For Rawls, Rousseau's general will is 'properly expressed in fundamental political laws concerning constitutional essentials and basic justice' (LHPP: 223). After the first standpoint of justice of the original position, this involves a further three stages in a four-stage sequence that defines appropriate points of view for considering different questions: second the justice of the constitution, third the justice of legislation, and fourth the application of rules by judges and legislators and the following of rules by citizens. Part two secondly involves taking up questions of political economy, how the economic and social system 'shapes the wants and aspirations that its citizens come to have', determining 'in part the sort of persons they want to be as well as the sort of persons they are' (TJ: 259/229). Since the design of institutions necessarily involves a view of the human good, and how institutions can realize this good, an ideal of the person should be embedded in the procedure through which principles for regulating those institutions are chosen. Institutions 'must not only be just but framed so as to encourage the virtue of justice in those who take part in them' (TJ: 261/231).¹⁰⁰ The condition of publicity, incorporated into the original position because of the role of moral principles in 'establishing the ties of community' (TJ: 582/510), guarantees that such a process can be carried through in a manner that respects moral autonomy. If one wanted to highlight Rawls's differences with Berlinian liberalism, one could say that the formation of citizens is to take place in accordance with a conception of their 'higher' political selves as free and equal citizens that they may freely endorse.

A view of social institutions as educating persons to a conception of themselves as citizens forms the context of Rawls's discussion of moral learning in part three, 'Ends'. This concerns the standpoint from which persons pursue with deliberative rationality

¹⁰⁰ Compare Rousseau: 'It is not enough to say to the citizens, *be good*; they must be taught to be so' (Rousseau 1997 [1755]: 142).

their conceptions of the good. It involves not Berlin's radical choice among plural values, but Josiah Royce's conception of the life plan that makes a conscious, unified moral person (TJ: 408/358). Rawls first postulates a basic principle of motivation, the Aristotelian principle, according to which, 'other things equal, human beings enjoy the exercise of their realized capacities (their innate or trained abilities), and this enjoyment increases the more the capacity is realized, or the greater its complexity' (TJ: 414/374). Rawls then turns to the problem of stability. The first part of the problem of stability concerns how persons motivated in such a way would acquire a sense of justice when living under the institutions of a well-ordered society. Rawls sees Rousseau's *Emile* as the instigator of a rationalist tradition – extended by Kant, sometimes by Mill, and by Piaget and Kohlberg – according to which right and justice are a natural outgrowth of the development of our social nature (TJ: 459/402-3). Here the family first plays the role of developing a morality of authority in children. Second, in the associational life of civil society persons develop a morality of association, 'the ideals of their station' in Bradley's sense (TJ: 470/429). Finally, in a more complex form of the morality of association – a morality of principles associated with an ideal of equal citizenship – persons develop a commitment to acting justly within, and advancing, just institutions.

The second part of the problem of stability concerns the question of the congruence of justice as fairness and goodness as rationality, the right and the good. The successful answer to this problem depends upon a well-ordered society achieving the good of community in the form of the social orchestra of a 'social union of social unions'. This permits the Aristotelian principle to have its 'wider effect' (TJ: 571/500) as the mutual cooperation of all brings forth the complementary latent powers of each member of society. In accordance with Rawls's 'Kantian interpretation' of justice as fairness, the desire to act justly is seen as congruent with the desire to express our nature as free moral persons. Somewhat ironically given later communitarian criticisms, for Rawls the Kantian interpretation is needed to make good on the claim that 'justice as fairness has a central place for the value of community' (TJ: 264/233):

Kant had little direct influence on Rawls's initial drafts of *A Theory of Justice* in the 1950s and 1960s (most of the first six chapters and chapter 8 on the sense of justice). The Kantian interpretation of justice as fairness . . . was written relatively late and was incorporated into *A Theory of Justice* primarily to show how justice as fairness is 'congruent' with the good (Freeman 2007a: 21-22).

Sandel (1996) describes the original position as a lens: looked through in one direction we have principles of justice, from the other a conception of the person. This is true, but Sandel overlooks the way in which the principles of justice chosen in the original

position are realized within the basic structure of the social system, this forming the pivot of Rawls's architectonic. The idea of the basic structure of a well-ordered society is needed along with the idea of the original position 'to complete other ideas and to order them into a perspicuous whole' (PL: 14-15 n.16). The persons of Part three who achieve social unity in a well-ordered society are not unencumbered selves, but rather integrated into society. They are the kind of persons that, given favourable conditions and appropriate institutions, we could be and therefore can reasonably desire to be.

Having set out the structure of TJ, and before examining in more detail the problem of stability, in the remainder of this section I seek to clarify three points that have led to misunderstandings of Rawls's view: (1) the relationship between deontology and the priority of the right, (2) the relationship between reflective equilibrium and the original position, and (3) the relationship between justice and legitimacy.

(1) Sandel's misconceptions result from eliding two issues: the distinction between deontological and teleological doctrines and the distinction between the priority of the right and the idea of the single rational good (Freeman 2007b: 45-75). Rawls says that justice as fairness is a deontological view by definition since it is not teleological: it does not have a conception of the single rational good. A deontological doctrine is '[o]ne that either does not specify the good independently from the right, or does not interpret the right as maximizing the good'. Justice as fairness, Rawls says, 'is a deontological theory in the second way', and therefore *not* in the way in which the intuitionism of H.A. Prichard and Ross identifies the good with acting according to the independently defined right. To characterise the rightness of institutions or acts independently of their consequences, or, we might add, basis in human nature, would be 'irrational, crazy' (TJ: 30/26). Principles of justice are chosen in accordance with a thin theory of the good, a basic rational motivation for social primary goods. The goodness as rationality of deliberative rationality is specified independently of the right, but while, thanks to the thin theory of the good at the heart of the right, persons can see principles of right as serving their good, principles of the right are prior to, in the sense of being regulative of, persons' pursuit of the good, rather than maximizing its realization. The way in which justice as fairness is deontological means that this issue is different from, but not unrelated to, the priority of the right. As in Kant so in Rawls, deontology and teleology are interrelated, although Rawls relies upon a social rather than a natural teleology. This is a teleological process of moral learning in accordance with the Aristotelian principle in which

each stage foreshadows in its teachings and explanations the conception of right and justice at which it aims and by reference to which we will later

recognize that the moral standards presented to us are justified . . . Thus moral education is education for autonomy (TJ: 515/452).

Principles of the right are prior to conceptions of the good in a teleological sense: they are the end within which and towards which conceptions of the good develop, not a dominant end but one where acting on principles of the right is reflectively endorsed as the culmination and completion of our rational good and can therefore be affirmed for its own sake.

The *congruence argument* purports to show that under the ideal conditions of a well-ordered society, the judgements that would be made from th[e] two ideal perspectives [the original position and deliberative rationality] coincide: reasonable principles judged and willed as rational from the common perspective of justice are also judged and willed as rational from each individual's point of view (Freeman 2007a: 266).

(2) The deductive choice of principles by the parties in the original position may seem hard to reconcile with the seemingly intuitive and inductive method of reflective equilibrium (Scanlon 2003a: 139). But as we saw in the last chapter, Rawls does not offer independent 'contract' and 'coherence' arguments (Lyons 1989): reflective equilibrium works through the original position. Reflective equilibrium is not a *method* in the sense of being an instrument applied to an external object, but rather a *process* of practical reasoning. While analysis of reflective equilibrium has tended to be restricted to questions of scope – narrow versus wide reflective equilibrium (Daniels 1996) – Scanlon (2003a) distinguishes between contrasting descriptive and deliberative conceptions of reflective equilibrium. What Scanlon calls the descriptive conception is discussed by Rawls in relation to the methodology of moral theory: 'one thinks of the moral theorist as an observer, so to speak, who seeks to set out the structure of other people's moral conceptions and attitudes' (CP: 288). Following Scanlon, Miller complains that this makes moral theory dependent on external facts: considered judgements are presented as 'akin to the raw data that might serve to ground a scientific theory' (Miller 2008: 41). We have already seen that Rawls takes considered judgements not as data but as reflecting general facts. Taking this further, an analogy between moral theory and scientific theory is apt if one sees Rawls's position as analogous to the double hermeneutic involved in a post-positivistic philosopher of science analysing the practice of scientists. Take the following characterisation of Quine's project, indebted to Dreben:

Insofar as Quine's argument that a science of the specifically mental is separable from 'acceptable' science is dependent on a prior characterization of 'acceptable' science, it is circular. But it is circular only by contrast with

linear arguments whose conclusions have only as much strength as they get from their premises. The argument is not simply circular, but something much more complicated. In arguing for the separability of acceptable science from a science of the specifically mental, Quine is doing scientific epistemology, and the scientific epistemologist has himself as part of his subject. So, in a sense, Quine's argument doesn't have two parts any more than it has one part. And, rather than by the metaphor of a two-dimensional circle, Quine's argument seems best described by analogy with Hegel's absolute becoming conscious of itself (Schuldenfrei 1972: 16).

This interpretation will require some unpacking later on, but it will guide us in considering the 'two parts' of Rawls's theory. Immediately after the passage quoted by Scanlon and Miller, Rawls says that, given the variation in people's conceptions of justice, we are best to proceed as theorists by studying the leading conceptions in the history of moral philosophy and by engaging in self-examination. We are interested in what moral conceptions people *would* affirm after systematic reflection on the leading historical conceptions, not in brute facts about what conceptions they *do* hold. This is science, but it is moral science. The moral theorist is a scientist in the manner of Kant's description of Rousseau as the 'Newton of the moral world'. Scanlon is right that a deliberative method of reflective equilibrium, one of figuring out what to believe about justice, is primary. But this is because while descriptive reflective equilibrium takes place from the standpoint of the moral theorist in establishing the original position, deliberative reflective equilibrium takes place from the standpoint of the deliberative rationality of citizens upon whose reflective endorsement the justification of the original position ultimately depends.

(3) This is related to a third point. Estlund (1996) has rightly emphasised that Rawls's egalitarian conception of justice – justice as fairness – remains unaltered in its fundamentals in the move from *TJ* to *PL*. However, Estlund's reason for this claim – that the latter is concerned with legitimacy whereas the former was concerned with justice – is unpersuasive.¹⁰¹ The real picture is more complex. *TJ* introduces a criterion of legitimacy into the choice of principles of justice through the incorporation of an ideal of persons as free and equal citizens into the original position. Justice must be justified justly. The incorporation of the demand of publicity is a synchronic constraint on acceptable principles that ensures that justification can be fully carried out, that justice as fairness can be realized in a manner consistent with the values it expresses. Stability is the diachronic aspect of publicity, the capacity of publicity to persist over time. It is the requirement that justice as fairness, unlike utilitarianism, be self-sustaining rather than self-effacing. As we will see, Rawls's conception of public reason in *PL* develops

¹⁰¹ Similar considerations apply to Wenar (2004). On the role of legitimacy in *PL*, see Kelleher (2008).

this notion of publicity (Larmore 2003), further responding to the demand for a just justification of justice.

6.4 Ideal Feasibility and Utopianism

I have outlined a blueprint for reading Rawls and, working within this, examined how Rawls constructs the architectonic of justice as fairness in *TJ*. I turn now to *PL*, Rawls's self-critique, or, as I will argue in 6.6, metacritique, of the idea of a well-ordered society of justice as fairness in *TJ*. Rawls argues that with hindsight this conception of a well-ordered society was 'unrealistic', where by this he means 'inconsistent with realizing its own principles under the best of foreseeable conditions' (PL: xix). Justice as fairness must become a political conception of justice because it is 'unrealistic', indeed 'utopian' (PL: 39), to expect a consensus on a comprehensive conception of justice. What is at fault is a 'serious problem internal to justice as fairness [arising] from the fact that the account of stability in Part III of *Theory* is not consistent with the view as a whole' (PL: xv-xvi).

In what sense might the argument for stability fail to live up to the requirement that political philosophy be 'realistically utopian'? One answer that immediately suggests itself is that *TJ* had a utopian comprehensive conception of a well-ordered society. As Rawls defines it, a comprehensive doctrine 'includes conceptions of what is of value in human life, as well as ideals of personal virtue and character, that are to inform much of our non-political conduct (in the limit our life as a whole)' (CP: 450). But if this is the case, one could be forgiven for a certain incredulity. As Barry argues, surely the 'whole point of *A Theory of Justice* was that it left people to form, revise, and pursue their own conception of the good', rather than imposing a perfectionist conception of the good upon them (Barry 1995b: 878). But on Rawls's conception, individuals must not only be motivated to do what justice requires, but this motivation must have appropriate psychological origins. In the original position, the parties do not just consent to principles of justice, but also 'consent to the arrangements necessary to make these principles effective in their conduct', including practices of moral education (TJ: 515/452). The stability argument is required to show that moral convictions are not the result of coercive indoctrination or ideological illusion.

For Barry, stability is a Hobbesian problem of compliance that has no place in an impartialist theory. Scanlon's conception of moral motivation as desiring to act upon principles that no one could reasonably reject provides Barry with a substantive criterion of rightness, and all that is necessary to ensure stability. Individuals who are not

motivated in this way will, by that fact alone, be unreasonable. Why, Barry asks, does Rawls reject Ross's idea of the purely conscientious act, the idea that individuals can be motivated to do the right thing for its own sake? Barry concludes, exasperated, that the only thing to do with the final chapter of *TJ* 'is to follow the course followed virtually unanimously by commentators . . . and forget about it' (Barry 1995b: 915 n.54). However, the doctrine of the purely conscientious act follows Moore in making the preference for justice external to individuals' rational good. Rawls argues, following the Hegelian J.N. Findlay, that this makes the desire for justice arbitrary (*TJ*: 477 n.15/418 n.15). What Rawls opposes is the idea of moral motivation as the 'desire to do what is right and just simply because it is right, *no other description being appropriate*' (*TJ*: 447/418, emphasis added). With the original position, though, we *do* have an appropriate description (Freeman 2007b: 148), as in understanding our sense of justice as the desire to act on principles that we could agree to in an initial situation of freedom and equality we understand the source of principles of justice in our reasonable nature.

As Freeman puts it, congruence between the right and the good 'requires that principles of justice, derived on grounds independent of given preferences, be within the reach of human capacities and be compatible with a human good that affirms our nature' (Freeman 2003b: 288). Picking up on Rawls's description of how justice as fairness is presented in two stages, Freeman says 'a conception of justice should be worked out beforehand by relying on independent moral considerations. Then the question of its stability is raised to test the *feasibility* of a just society conceived along the lines of this conception' (Freeman 2007b: 144-45, emphasis added). But what Freeman calls Rawls's Kantian congruence argument, Rawls came to see as inconsistent with reasonable pluralism, since

we cannot expect large numbers of people in a well-ordered society to be motivated to comply with standards of justice for the Kantian reason that they realize their nature as free and equal rational moral beings and are thereby morally autonomous (Freeman 2007b: 170).

However, accepting that Rawls's dissatisfaction with *TJ* results from the perceived failure of the Kantian congruence argument, ambiguities remain.¹⁰² We might ask first why the need to revise the second-stage argument for stability leads the independently derived conception of justice as fairness to stand condemned as a utopian comprehensive

¹⁰² Weithman (2007) argues against Freeman that Rawls's dissatisfaction with part three of *TJ* extended beyond the Kantian congruence argument to include the difficulties reasonable pluralism creates for Rawls's conception of a well-ordered society. However, I can see no grounds for thinking that Rawls revises his basic conception of a well-ordered society in *PL*; that society can be a social orchestra is not questioned (*PL*: 204, 321).

doctrine. Rawls says in *TJ* that ‘if justice as fairness succeeds reasonably well, a next step would be to study the more general view suggested by the name “rightness as fairness”’ (*TJ*: 17/15). But with hindsight Rawls comments that from this ‘the reader might reasonably conclude . . . that justice as fairness is set out as part of a comprehensive view that may be developed later were success to invite’ (*CP*: 489). Barry argues that ‘only a very obtuse reader’ would conclude any such thing (Barry 1995b: 878). We might ask, secondly, why Rawls is concerned that *TJ* is inconsistent with realizing its own principles under the best foreseeable circumstances. For *TJ* ‘supposes that the well-ordered society of justice as fairness is possible and somehow comes about. It then asks whether the society is stable’ (*PL*: xlvi, emphases added). It does not ask how it might be realized under existing conditions.

In fact, Barry is right to argue that justice as fairness is *not* a comprehensive doctrine. What is really at issue for Rawls is the question of scope, ‘the range of subjects to which a conception applies and the content a wider range requires’ (*PL*: 13). Rawls defines a moral conception as general ‘if it applies to a wide range of subjects, and in the limit to all subjects universally’, observing that ‘many religious and philosophical doctrines aspire to be both general and comprehensive’ (*PL*: 13). A political conception, by contrast, ‘tries to elaborate a reasonable conception for the basic structure alone and involves, so far as possible, no wider commitments to any other doctrine’ (*PL*: 13). If this is the key point it shows why Rawls cannot follow Barry in endorsing Scanlon’s notion of rightness as reasonable rejectability, since this follows Rawls’s suggestion of an account of rightness as fairness (Freeman 1991: 281). For Scanlon, the domain of ‘what we owe to each other’, if not covering the whole of morality, expands beyond the political domain of justice. While it may not be ‘utterly capricious’ (*TJ*: 478/418) to make principles derived in this way regulative of the basic structure neither, in drawing on Scanlon, does Barry have a principled basis for limiting the domain of justice to the basic structure. Scanlon’s contractualism does not seek to provide a public basis of justification consistent with the demands of full publicity and stability (Freeman 2007c: 36).

In one sense, then, the move to *PL* is needed to remove an ‘ambiguity’ (*PL*: xix). *TJ* did *not* make utopian assumptions about the good, but in not taking account of the general fact of the reasonable pluralism of our social world, justice as fairness is utopian in the other sense of the word, for it can find *no place* within our social world. *TJ* understood being ‘realistic’ in terms of ‘ideal feasibility’: human nature and the natural facts of our world are such that a well-ordered society is feasible. From the perspective of *PL*, rather than saying that *TJ* failed to be realistic, it would be more accurate to say that Rawls revises his conception of what philosophy can accomplish, of what realistic

utopianism consists in. If justice as fairness is to find a place in our social world, then the formation of citizens must be congruent with the stages through which persons are formed in familial and associational life. We can no longer just *suppose* that a well-ordered society comes about and *then* consider moral learning according to a process of moral development from birth.

In a rhetorically nuanced passage of *TJ*, Rawls says:

If men's conceptions of justice finally turn out to differ, the ways in which they do so is a matter of first importance. Of course we cannot know how these conceptions vary, or even whether they do, until we have a better account of their structure. And this we now lack, even in the case of one man, or homogenous group of men. If we can characterize one (educated) person's sense of justice, we might have a good beginning toward a theory of justice. We may suppose that everyone has in himself the whole form of a moral conception. So for the purposes of this book, the views of the reader and the author are the only ones that count (TJ: 50/44).

Now, according to Rawls we turn to moral philosophy when we are divided within ourselves. We will not be able to make progress in moral philosophy, however, without first making progress in moral theory, analysing the structure of moral conceptions. *TJ* is, moreover, presented as a propaedeutic to moral theory. There is a suggestion in the above passage of Rawls conducting an internal dialogue, or self-analysis, uncertain whether what he is describing is simply his own sense of justice or one that his readers can, and will, reflectively endorse as also their own. Rawls assesses *TJ* according to what a *reasonable reader*, not a professional philosopher, might conclude, and what they might reasonably conclude about the *direction in which the argument could be developed*. Therefore as well as removing ambiguities that made justice as fairness liable to misinterpretation, *PL* is needed to consider how a shared sense of justice can nevertheless be compatible with a plurality of reasonable comprehensive doctrines.

The problem is that in *TJ* Rawls assumed a homogenous sense of justice:

The original position is so characterized that unanimity is possible; the deliberations of any one person are typical of all. Moreover the same will hold for the considered judgements of the citizens of a well-ordered society effectively regulated by the principle of justice. Everyone has a similar sense of justice and in this respect a well-ordered society is homogenous (TJ: 263/232).

While avoiding many of the most serious faults of utilitarianism, justice as fairness will be self-effacing if, in willing the end of a well-ordered society, persons will have adopted a general moral point of view. Even if this is not used to elaborate a fully comprehensive doctrine, a general moral point of view leaves open the possibility that

willing a well-ordered society as an end in itself is conditional upon the comprehensive doctrine a person affirms. What makes justice as fairness utopian in *TJ* is that it was ‘general in range’ and this *left open the possibility* of its being ‘comprehensive in scope’ (CP: 482). Justice as fairness was inconsistent with itself since the first two parts assumed that the range of its principles was limited to the basic structure, whereas part three suggests an inquiry into human nature as such, independent of existing social and institutional forms. It would fail to meet the criteria that

a theory should present a description of an ideally just state of affairs, a conception of a well-ordered society such that the aspiration to realize this state of affairs, and to maintain it in being, answers to our good and is continuous with our natural sentiments. A perfectly just society should be part of an ideal that rational human beings could desire more than anything else once they had full knowledge and experience of what it was (TJ: 477/417-8).

What is required is a radicalisation of the principle of publicity: principles must not only be public, and based on publicly acceptable general facts, but, in addition, full publicity requires, as Rawls somewhat enigmatically puts it, that ‘justification includes everything that we would say – you and I – when we set up justice as fairness and reflect why we proceed in one way rather than another’ (PL: 67). Justice as fairness must not only be based on the generally shared public beliefs that are ascribed to the parties in the original position, but the very procedure of justification itself must be publicly justifiable if nothing is to be hidden from view.¹⁰³

6.5 Real Feasibility and the Search for Stability for the Right Reasons

In addressing the problem of stability in *PL*, the first question posed is the same as in *TJ*, ‘whether people who grow up under just institutions acquire a normally sufficient sense of justice so they generally comply with those institutions’ (PL: 141). It is also answered in the same way, through a conception of moral psychology, although this becomes a purely philosophical rather than a scientific conception (PL: 86). However, in place of the problem of Kantian congruence, the second question becomes whether ‘in view of the general facts that characterize a democracy’s public political culture, and in particular the fact of reasonable pluralism, the political conception can be the focus of an overlapping consensus’ (PL: 141).

¹⁰³ Contrasting interpretations of Rawls’s conception of stability are found in Copp (1996); Huemer (1996); Krasnoff (1998); Mendus (1999); Wingenbach (1999); Midgaard (2003). None understand stability in the sense currently at issue.

As already discussed in relation to ‘The Sense of Justice’, stability for Rawls was never a purely practical matter. In *TJ*, Rawls contrasted the stability of social cooperation under justice as fairness with Hobbesian and utilitarian conceptions. In place of a Hobbesian sovereign, Rawls based the stability of social cooperation on the very different basis of a collective sense of justice. A conception of justice is stable ‘when the public recognition of its realization by the social system tends to bring about the corresponding sense of justice’ (TJ: 177/154). We can therefore compare the relative stability of justice as fairness and utilitarianism, and in this respect Rawls argues that the notion of reciprocity in justice as fairness has advantages over the utilitarian conception of sympathetic identification with others. The parties in the original position would favour justice as fairness because reciprocity protects and supports the most important social primary good, that of self-respect, and ‘[s]elf-respect is reciprocally self-supporting’ (TJ: 179/156). Whereas under a social system regulated by justice as fairness, identification with the good of others may become a strong element of our own good, whatever our social position, a utilitarian society is likely to break the strains of commitment by asking the worst off to altruistically sacrifice their own good in the interests of maximising overall utility (TJ: 501/438).

The stability of the social system as Rawls describes it does not depend upon an unchanging equilibrium. Rather, an equilibrium in a social system is stable when ‘however institutions are changed, they still remain just or approximately so, as adjustments are made in view of new social circumstances’ (TJ: 458/401). It is in describing how civil disobedience can introduce stability into a well-ordered society, which does not rely upon an unchanging consensus, that Rawls first used the expression ‘overlapping consensus’ (TJ: 388/340). In *PL*, overlapping consensus takes on the different role of resolving the problem of congruence.

In his political turn, Rawls moves from moral philosophy to political philosophy, the latter being concerned with ‘practical political possibilities’ (CP: 447). While a moral conception ‘may condemn the world and human nature as too corrupt to be moved by its precepts and ideals’, a political conception must ‘fall under the art of the possible’. (CP: 486). The diversity of comprehensive religious, philosophical and moral doctrines in modern democratic societies is a permanent feature of their public political culture, such that a consensus on a comprehensive doctrine could only be maintained through the oppressive use of state power. It looks here as if Rawls is concerned with ‘real feasibility’, compromising with existing facts about the limits of moral motivation. But Rawls is instead concerned with the question of how ‘those who grow up under basic institutions that are just – *institutions that justice as fairness itself enjoins* – acquire a reasoned and informed allegiance to those institutions sufficient to render the institutions

stable' (CP: 487, emphasis added). We must consider 'the effects of the social conditions required by a conception of political justice on the acceptance of that conception itself' (CP: 414).

Speaking against the idea that the second stage is concerned with real feasibility is the fact that Rawls does not say that justice as fairness has two categorically distinct stages, but that it is '*best presented* in two stages' (CP: 486, emphasis added).

The problem of stability is not that of bringing others who reject a conception to share it, or to act in accordance with it, by workable sanctions, if necessary, as if the task were to find ways to impose the conception once we are convinced it is sound. Rather, justice as fairness is not reasonable *in the first place* unless in a suitable way it can win its support by addressing each citizen's reason, as explained *within its own framework* (PL: 143-44, emphases added; see also CP: 488).

Just as the two stages of the argument in *TJ* take place within the original position, so in the two stage architectonic of *PL*, 'both parts are carried out within the same framework and subject to the same conditions embedded in the original position as a device of representation' (PL: 143 n.10).

But if Rawls is not concerned with real feasibility, in the initial stages of his political turn the compatibility of a well-ordered society with reasonable pluralism is not clear. What Rawls refers to at this stage as the 'burdens of reason' (CP: 478) are reminiscent of Rousseau's 'yoke of reason' (Rousseau 1979 [1762]: 315). They are not integrated into Rawls's account of the fact of pluralism and play no positive role as the basis for the endorsement of a conception of justice. Rawls is still searching for *stability for the right reasons*, that is, stability on the basis of reasons beyond the scope of the formerly general point of view of the original position, but reasons that can nevertheless be reasonably explained within the framework of justice as fairness.

6.6 Realizability and Reasonable Pluralism

Rawls remarks that the need for a new family of ideas in *PL* may seem surprising:

the problem of stability has played very little role in the history of moral philosophy, so it may seem odd that an inconsistency of this kind should turn out to force such extensive revisions. Yet the problem of stability is fundamental to political philosophy and an inconsistency there is bound to require readjustments (PL: xix).

In *PL*, though, a political conception of justice remains a moral conception, one restricted to the domain of the political. A clue to Rawls's meaning is provided by his remark that for Hegel, 'many of the traditional ambitions of moral philosophy are to be brought within those of political philosophy, as Hegel understands it' (LHMP: 330). It is in a post-Kantian reading of Hegel that Rawls finds a model for a moral conception of political philosophy, and it is with reference to Hegel that the problem of stability is best understood. A moral-political conception of justice must not only be self-sustaining, but part of a self-reproducing practice of social cooperation.

Why should a theory reflect upon its capacity to fulfil a social role? Why not simply be concerned, like Barry, with whether it is true, or right?

The fact that Rawls's project proceeds in this self-reflective way, envisaging a role for itself in the maintenance and reproduction of a just society, is a deep fact about it, one that is to a large degree rooted in the modern idealist tradition in political philosophy, running from Rousseau, say, through Kant, Fichte and Hegel (Sensat 2007: 1).

Such 'anticipation of the conditions of its possible application' is for Horkheimer what distinguishes critical from traditional theory (TP: 1-2). Far from faulting Rawls for his philosophical modesty in *PL*, we might more properly ask with Patrick Neal 'Does he mean what he says?'

The "practical turn", which most commentators have taken to be a retreat from the ambition embodied originally in *A Theory of Justice*, appears instead as the mark of an even bolder enterprise than the original; an attempt not merely to theorize justice, but to realize it (Neal 1997: 130).

I have suggested that *PL* should be understood as a metacritique of *TJ*. The idea of metacritique begins with Hegel's critique of Kant in the *Phenomenology of Spirit*, is radicalised by Marx, and provides the motive force behind Frankfurt School critical theory up until Habermas's self-critique of *Knowledge and Human Interests*. Metacritique represents a radicalisation of critique:

In the process of critique, the object criticized is subjected to examination, but the critical canon [the presuppositions of critique] remains in the dark. There is no question of seeing the suspicion of error displaced from the object criticised onto the critical subject. The subject engaged in critique does not reflect on its relation to the object criticised . . . What is revealed here cannot be an ultimate foundation, posited in an absolute manner as in metaphysics. This experience reveals metacritique as a radicalisation of critique, and the movement associated with the 'meta' is only radical as long as it resolutely refuses any such absolute position (Kortian 1980: 29-30).

The light of critique must reach behind the idealizing mirror that justice as fairness holds up to our sense of justice to reflect upon the possibility and presuppositions of its offering this reflection. As Cohen recognises, the aspiration of the Hegelian Marxism that critical theory inherits is to be doubly reflexive in its relation to practice. ‘Gazing into the mirror of the historical context of its own genesis, it also discovers the addressee who can be spurred on to a liberating practice by means of the critical insights that the theory provides’ (TaJ: 282; see also TP: 1-2). It is this that explains Rawls’s idea that we must ‘apply the principle of toleration to philosophy itself’ (CP: 395). Justice as fairness seeks to recognise doctrines of toleration as among the subjective circumstances of justice and then spell out the implications of the principle of toleration for philosophy itself (CP: 412). In relating to its object in this way, a political conception of justice can hope to achieve doctrinal autonomy (PL: 98-99). It is required to do so by a radicalisation of the demand of publicity that the value of autonomy brings with it.

The distinction between the fact of pluralism and the fact of *reasonable* pluralism is first drawn by Joshua Cohen. As Cohen puts it, ‘in a Hegelian *Doppelsatz*: We need to accommodate the ideal to the real because the real manifests the ideal’ (Cohen 1993: 288).¹⁰⁴ The relevant question is ‘what sorts of constraints on realizability [of norms of justice] are constitutive of ideal justice’ (1993: 271). Or as Rawls puts it,

in framing a political conception of justice so it can gain an overlapping consensus, we are not bending it to existing unreason, but to the fact of reasonable pluralism, itself the outcome of the free exercise of free human reason under conditions of liberty (PL: 144).

An overlapping consensus is not the best possible form of consensus, but is ‘sufficient for the most reasonable basis of social unity available to us’ (PL: 149). We are looking for the basis of toleration, given the fact that reasonable pluralism is the inevitable outcome of free institutions, the ‘deepest and most reasonable basis of social unity available to us in a modern democracy’ (PL: 391).

Rawls’s answer to this problem makes use of the concept of the reasonable. As a virtue of persons, the reasonable has two aspects. The first aspect is that persons are reasonable when ‘they are ready to propose principles and standards as fair terms of cooperation and to abide by them willingly, given the assurance that others will likewise do so’ (PL: 49). The second virtue of the reasonable person is ‘the willingness 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 in a constitutional regime’

¹⁰⁴ On Hegel’s *Doppelsatz*, or ‘double dictum’, that ‘What is rational is actual; And what is actual is rational’, see Hardimon (1994). Cf. Stern (2006).

(PL: 54). The burdens of judgment are defined as ‘the many hazards involved in the correct (and conscientious) exercise of our powers of reason and judgment in the ordinary course of political life’ (PL: 56). Rawls says that

being reasonable is not an epistemological idea (though it has epistemological elements). Rather, it is part of a political ideal of democratic citizenship that includes the idea of public reason. The content of this ideal includes what free and equal citizens as reasonable can require of each other with respect to their reasonable comprehensive views (PL: 62).

Recognising the consequences of the burdens of judgement involves accepting a reasonably realistic ideal of a well-ordered society. Reasonable pluralism is affirmed as constitutive of the most reasonable form of social unity available to us.

This chapter has reconstructed the development of Rawls’s project up until *PL* within the framework of a self-grounding constructivist blueprint for reading Rawls that emphasises his increasing appropriation of ideas from Rousseau and Hegel and awareness of the realizability relation between democratic ideals and general facts. However, it has raised a number of questions that will require further discussion in the next chapter.

7. Overlapping Consensus and the Unrealizability of Reconciliation

In this chapter, an examination of the Rawls-Habermas debate will provide an opportunity both to analyse Rawls's most well-developed account of the role of an overlapping consensus and to begin to uncover the force of Habermas's critique of Rawls. Habermas's initial critique of Rawls, 'Reconciliation through the Public Use of Reason' (HR 1), is divided into three sections, and Rawls's 'Reply to Habermas' follows this structure (Table 2).

<u>Habermas</u>	<u>Rawls</u>
1. The Design of the Original Position	1. Two Main Differences a) Political vs. metaphysical doctrines b) Original position vs. ideal discourse situation
2. The Fact of Pluralism and the Idea of an Overlapping Consensus	2. Overlapping Consensus and Justification
3. Private and Public Autonomy	3. Liberties of the Moderns Versus the Will of the People
	4. The Roots of the Liberties
	5. Procedural Versus Substantive Justice

Table 2. The Rawls-Habermas Debate

In Part I, I criticised the way that Habermas clouds the issues at stake in the debate by combining criticisms from the perspective of his moral theory with criticisms from the perspective of his political theory. A meta-framework is needed in relation to which to situate their respective approaches, and I have argued that this is provided by the framework of a post-Kantian interpretation of Hegel. The subsequent analysis has lent support to the claim that the role of the original position is misunderstood when it is seen as representing a standpoint of individual moral judgement divorced from a structure of background institutions, and I shall not focus on this aspect of Habermas's critique.

More penetrating are Habermas's criticisms regarding the role of an overlapping consensus. Here he makes a set of interrelated points about the relationship between the two stages of Rawls's theory. The first concerns the role Habermas takes reflective equilibrium to play at these two stages. Habermas here returns to themes he had already addressed in an analysis of Rawls in *BFN* that served as a springboard for the

presentation of his own approach. This point is best left to Chapter 9 in which I discuss *BFN*. Habermas's more immediate concern in his debate with Rawls is that an overlapping consensus serves merely the functional role of ensuring social stability as opposed to being a confirmation of the justifiability of justice as fairness. This is the first point to be addressed in this chapter, and, building on the analysis of the last chapter, we will see that Habermas's critique is weakened by abstracting Rawls's conception of justice from Rawls's background social theory. This is confirmed by Rawls's reply to Habermas, in which Rawls gives his most detailed account of the role of an overlapping consensus within a process of justification.

Habermas's criticisms are more successful, however, when he examines the relationship between citizens' reasonable comprehensive doctrines and a moral standpoint, which, Habermas argues, must claim philosophical truth. Dissatisfied by Rawls's reply to this point, Habermas returns to this issue in “Reasonable” versus “True,” or the Morality of Worldviews’ (HR 2). Although Habermas points towards fundamental problems with Rawls's approach, and the way it relates private and public autonomy, uncovering them requires first recognising the way in which Rawls's Rousseauian conception of public reason differs from Habermas's Kantian conception of the public *use* of reason. I will argue that Habermas's criticisms are ultimately best understood in relation to the problem of self-grounding. Habermas's strictly procedural approach separates a conception of procedural reason from a substantive ideal of a well-ordered society and thereby avoids the aporias of a utopian promise of reconciliation to our social world. Constitutional democracy is reconceived as an ongoing project lacking recourse to external guarantees of stability and reproducing itself from its own resources. This contrast with Rawls will be more fully elaborated in the remainder of this thesis, but this argument begins in the final two sections of this chapter where I establish initial grounds for arguing that Habermas offers the best avenue for pursuing post-Rawlsian political philosophy.

7.1 Overlapping Consensus and Self-grounding Justification

For Habermas, the reason why Rawls can only present an overlapping consensus as confirming the utility of a political conception of justice in ensuring stable social cooperation is to be found in the way that Rawls constructs the moral point of view of the original position. After the justification of principles of justice, the veil of ignorance is gradually lifted, but

If we are to ensure that no discrepancies arise, we must construct the original position already with knowledge and even foresight, of all the normative contents that could potentially nourish the shared self-understanding of free and equal citizens in the future. In other words, the theoretician himself would have to shoulder the burden of anticipating at least parts of the information of which he previously relieved the parties in the original position! (HR 1: 58).

In an earlier discussion of Rawls's theory in comparison with his discourse theory of morality, Habermas put the point in a less conciliatory tone: Rawls is said to view his substantive conclusions 'not as the *contribution* of a participant in argumentation to a process of discursive will formation . . . but as the outcome of a "theory of justice," which he as an expert is qualified to construct' (MCCA: 66).

Habermas argues that the test of the acceptability of justice as fairness to citizens with a plurality of comprehensive doctrines in *PL* is of a qualitatively different kind to the test of a well-ordered society's capacity to be self-sustaining in *TJ*. Rawls's 'methodological parallel is problematic because the test cannot be undertaken in an immanent manner in the case of acceptability; it is no longer a move within the theory' (HR 1: 61). Philosophical theory cannot, and should therefore cease its attempt to, anticipate the outcome of real discourses. In seeking to justify substantive principles rather than simply the necessity of adopting the moral point of view, an overlapping consensus serves only the pragmatic goal of ensuring the acceptance of substantive conclusions, rather than embodying a procedure of discourse about substantive principles that reflexively confirms the justified acceptability of the moral point of view. Rawls seems deliberately to elide the distinction between justification and stability:

Because Rawls situates the 'question of stability' in the foreground, the overlapping consensus merely expresses the functional contribution that the theory of justice can make to the peaceful institutionalization of social cooperation; but in this the intrinsic value of a *justified* theory must already be presupposed. From this functionalist perspective, the question of whether the theory can meet with the public agreement – that is, from the perspective of different worldviews in the forum of the public use of reason – would lose an epistemic meaning essential to the theory itself. The overlapping consensus would then be merely an index of the utility, and no longer a confirmation of the correctness of the theory; it would no longer be of interest from the point of view of acceptability, and hence of validity, but only from that of acceptance, that is, of securing social stability. When [Rawls] calls his conception of justice 'political' his intention appears rather to be to collapse the distinction between its justified acceptability and its actual acceptance (HR 1: 62).¹⁰⁵

¹⁰⁵ As I noted in Chapter 2, there are superficial structural similarities between Habermas's critique and Cohen's. We will be in a position to note the depth of the differences in Chapter 10.

The fact that as the veil of ignorance is lifted, and Rawls's citizens take on flesh and blood, they find themselves subject to already anticipated and institutionalized principles means that they experience these principles as beyond their control. They cannot regard the constitution as an ongoing radical democratic project:

From the perspective of the theory of justice, the act of founding the democratic constitution cannot be repeated under the institutional conditions of an already constituted just society, and the process of realizing the system of basic rights cannot be assured on an ongoing basis. It is not possible for the citizens to experience this process as open and incomplete. Citizens cannot reignite the radical democratic embers of the original position in the civic life of their society, for from their perspective all the *essential* discourses of legitimation have already taken place within the theory; and they find the results of the theory already sedimented in the constitution (HR 1: 69-70).

Before we can assess these charges we will need to examine further the role of an overlapping consensus. In his reply to Habermas, Rawls sets out how a reasonable overlapping consensus plays a role in public justification alongside the ideas of stability for the right reasons and legitimacy (PL: 394). Public justification is the culmination of three stages of justification. In the first stage, *pro tanto* justification involves a justification of a political conception, presented as freestanding, solely according to political values. In a second stage, full justification involves citizens as individual members of civil society embedding the 'module' of a political conception into their comprehensive doctrine, perceiving that conception as reasonable or true, depending upon the comprehensive doctrine they hold. In a final stage, public justification by civil society happens when all reasonable citizens carry out justification of the shared political conception by embedding it in their reasonable comprehensive doctrines, mutually acknowledging their fellow citizens as doing likewise. At this final stage of justification, a reasonable overlapping consensus is formed. The hope is that citizens will then judge that political values 'outweigh or are normally . . . ordered prior' to nonpolitical values affirmed by comprehensive doctrines (PL: 392). Since a political conception of justice is not a compromise between already existing reasonable comprehensive views, it achieves 'stability for the right reasons' (PL: 390). It will have been shown that democratic legitimacy – the exercise of political power in accordance with a constitution that can be endorsed by free and equal citizens – can be fulfilled despite the existence of reasonable pluralism.

In public justification, a public political conception of justice is held by citizens in wide and general reflective equilibrium (PL: 388). This reflective equilibrium is wide because of the wide-ranging reflections of citizens who have taken into account the main

conceptions of political justice in our philosophical tradition. It is general since when all citizens have achieved wide reflective equilibrium their several reflections will have jointly achieved a well-ordered society in which all affirm the same public political conception of justice, one that Rawls argues is fully intersubjective (PL: 385 n. 16). Wide and general reflective equilibrium with respect to public justification gives ‘the best justification of the political conception that we can have at any given time’ (PL: 388). A political conception is therefore self-grounding, or self-standing. Its principles, and the normative grounds according to which they are *pro tanto* justified, are, through their capacity for full publicity, publicly seen to be grounded solely in the public values of the domain of the political.

Habermas, though, remains dissatisfied with the role of an overlapping consensus, seeing it as decisively weakening a Kantian strategy of justification by allowing the intrusion of Hobbesian elements. For Rawls, ‘Agreement in conclusions *results* from premises rooted in different outlooks’ (HR 2: 84). Political justice and private morality are connected, but the nature of this relation remains publicly inaccessible, depriving the political sphere of an independent validity. Citizens can only mutually observe that others endorse a political conception, the reasons for that endorsement being veiled by comprehensive doctrines that an observer’s perspective cannot penetrate. An overlapping consensus is therefore a ‘private use of reason with public-political intent’ (Rainer Forst, quoted in HR 2: 91). For Habermas, on the other hand, ‘even outside the political domain, arguments call for a public use of reason’ (HR 2: 86). The transition from an ‘ethical’ to a ‘moral’ use of practical reason, which corresponds to the transition from a comprehensive doctrine to a political conception, is ‘not a matter of a shift in perspective from internal monological thought to public discourse but of a transformation of the problem at issue; what changes is the role in which other subjects are encountered’ (JA: 15). Citizens cannot grasp the reasonableness of principles if it is not open to them to adopt a third standpoint besides that of an observer or participant, a first-person plural perspective from which all citizens support a conception of justice for the *same* reasons.

However, in switching his perspective from that of his moral theory to that of his political theory, Habermas presupposes the validity of the conclusion to his earlier argument that the moral point of view should be reconstructed in accordance with the procedural requirements of the public use of reason. Making the same mistake as many interpreters of Rawls in seeing *PL* as a wholly different project to *TJ*, he does not acknowledge the fact that a conception of justice justified through the original position is retained in *PL*, providing the substantive public framework that makes an overlapping consensus possible. Rawls *does* claim to have assembled the original position so that the

principles of justice as fairness ‘nourish the shared self-understanding of free and equal citizens’ in making possible the realization of a well-ordered society. While it is therefore true that citizens are ‘denied’, or, better, are not seen as possessing, the “moral point of view” from which they could develop and justify a political conception *in joint public deliberation*’ (HR 2: 84), it does not follow that an overlapping consensus is a moral compromise that lacks a regulative moral point of view: this is provided by the original position. Rawls does not presume the truth of comprehensive doctrines, but rather seeks to remain agnostic, since justice as fairness seeks to make a constructive contribution to the formation of an overlapping consensus. It nevertheless remains the case that, if successful, *a posteriori* a political conception of justice will be regarded by citizens who are part of an overlapping consensus as of a different order of validity to comprehensive doctrines, analogous to Habermas’s distinction between moral and ethical validity. The real question is not, then, ‘whether Rawls can account for the possibility of an overlapping consensus without tacitly assuming . . . a third perspective from which “we”, the citizens, can publicly examine in common what is in the equal interest of everybody’ (HR 2: 87), but whether he can do so without tacitly assuming the authority of the third perspective of the original position which claims to identify what is in the equal interest of everybody.

7.2 Reconciliation by Public Reason

Habermas fails to accurately characterise Rawls’s approach to the problem of self-grounding because he interprets Rawls according to a Kantian framework, following Kohlberg in seeing the original position as a moral point of view for guiding individual moral action. Only very belatedly does Habermas acknowledge, following Rainer Forst, that ‘Rawls salvages a valuable insight of Hegel’s critique of Kant; moral norms must be *internally* related to the life-plans and lifestyles of affected persons in a way they can grasp for themselves’ (HR 2: 100). The analysis of Rawls’s project in the previous chapter provides grounds for interpreting Rawls as seeking to integrate this Hegelian insight, and therefore a Hegelian rather than a Hobbesian concern with stability, into a Kantian justificatory framework. It is significant that from the perspective of his own self-described Kantian Republicanism, Habermas entitles his initial critique of Rawls ‘Reconciliation through the Public Use of Reason’.¹⁰⁶ But Rawls, so far as I can see, nowhere uses this phrase. By focusing first on the concept of public reason we will be in

¹⁰⁶ Habermas explicitly equates Rawls’s public reason with Kant’s public use of reason in BR: 135; BNR: 119.

a position to then see the role of public reason in contributing to the Hegelian end of reconciliation to our social world.

Before his development of the idea of public reason, Rawls did refer to the ‘free public use of our reason’ in relation to free speech (PL: 348), acknowledging that the term is inspired by Kant.¹⁰⁷ But in other contexts Rawls refers to the ‘*use of* public reason’ (PL: 54, 240, emphasis added). Public reason embodies a Rousseauian general will:

Public reason is characteristic of a democratic people: it is the reason of its citizens, of those sharing the status of equal citizenship. The subject of their reason is the good of the public: what the political conception of justice requires of society’s basic structure of institutions, and of the purposes and ends they are to serve (PL: 213).

We must be careful not to conflate Rawls’s social ontology with Habermas’s. Rawls cautions that ‘[t]he public reason of political liberalism may be confused with Habermas’s public sphere but they are not the same’ (PL: 382 n.13). Public reason applies to the public political forum: to the discourses of judges, with the Supreme Court as the exemplar of public reason; the discourses of government officials; the discourses of candidates for public office; and to citizens when voting on constitutional essentials and matters of basic justice. The public political forum is distinct from the background culture of civil society, which is the sphere of nonpublic (not private) reason, and to which public reason does not apply. Rawls is ‘concerned with reason, not simply with discourse’ (PL: 220). It is the background culture which Rawls equates with Habermas’s public sphere (PL: 382 n.13; CP: 582-83 n.28).¹⁰⁸ Rawls emphasises these points in response to Seyla Benhabib (1994: 36-37), but it is not, as Benhabib claims, that Rawls locates the public sphere in the state rather than in civil society, but that Rawls’s social ontology has no place for a public sphere in Habermas’s sense.¹⁰⁹

What is the significance of public reason for the idea of reconciliation to our social world? Rawls refers not to reconciliation through the public *use* of reason, but instead to ‘reconciliation through public reason’ (CP: 395), and more often to ‘reconciliation *by* public reason’ (PL: lx, 157; CP: 612, emphasis added).¹¹⁰ This difference is far from simply semantic. One of the limits to reconciliation by public reason is that

¹⁰⁷ See also PL: 336, 346, 348, 354, 355. In attributing the phrase to Kant (1991 [1784]), Rawls does not mean to endorse Kant’s distinction between the free public use of reason and its private use (PL: 213 n.2, 296 n.13).

¹⁰⁸ Rawls later includes a nonpublic political culture comprising media of all kinds (CP: 576 n.13).

¹⁰⁹ For a defence of Rawls’s position, see Charney (1998).

¹¹⁰ Rawls also refers to ‘reconciliation by free public reason’ (CP: 440).

‘comprehensive doctrines are, politically speaking, unreconcilable’ (PL: 91). While the differences between citizens that result from their differing comprehensive doctrines cannot be reconciled, we can be reconciled to conflicts produced by different status, class position, and occupation, or from ethnicity, gender and race, when we accept reasonable principles of justice and know that our social institutions conform to them. The political relation in a constitutional democratic regime can still be one of civic friendship (CP: 579). In reconciliation by public reason we become reconciled to our social world, affirming reasonable pluralism as constitutive of our freedom rather than a limitation upon it. Public reason for Rawls has an ontological dimension as much as an epistemological one. Rawls accepts ‘the Kantian (not Kant’s) view that what we affirm on the basis of our free and informed reason and reflection is affirmed freely’. But Rawls interprets this in a post-Kantian Hegelian sense

Freedom at the deepest level calls upon the freedom of reason, both theoretical and practical as expressed in what we say and do. Limits on freedom are at bottom limits on our reason: on its development and education, its knowledge and information and on the scope of the actions in which it can be expressed, and therefore *our freedom depends on the nature of the surrounding institutional and social context* (PL: 222-23 n.9, emphasis added).

As well as being a relation of free and equal citizens who exercise ultimate political power as a collective body, the fundamental relation of citizenship is ‘a relation of citizens within the basic structure of society, a structure we enter only by birth and exit only by death’ (CP: 577). Rawls’s conception of an overlapping consensus is not a private use of reason with public political intent. Rather, a substantive framework of public reason regulating the relation of citizens within the basic structure can reconcile citizens to the irreconcilable differences between comprehensive doctrines by showing how such doctrines nevertheless provide the nonpublic reasons on the basis of which a political conception of justice can be endorsed.

Habermas would be right to reply, though, that even if this makes Rawls’s view intelligible, it comes at a high, indeed ultimately unacceptable, price. Rawls’s social ontology, and its associated account of the motivational bases of citizens’ political deliberations, is, in Habermas’s terms, communitarian. For Rawls, justice ‘requires a relation, not to persons, but to institutions that embody principles of justice’ (Alejandro 1993: 88), or, more precisely, a relationship between persons mediated by just institutions. Habermas misrepresents Rawls’s position when he equates Rawls’s conception of reasonableness with his own conception of the demands of postmetaphysical thinking. Holders of comprehensive doctrines do not regard their

comprehensive doctrines in Habermas's reflexive manner, relativising their own pursuit of such doctrines to that of others. An intersubjective relationship is formed only via the mediation of a political conception of justice for the basic structure, and it is only on this basis that citizens can adopt such a reflexive attitude.

Comprehensive doctrines are transcendent in the sense that their content is formed independently of an intersubjective context of socialization and in that they cover a complete scheme of value. Indeed, one might question whether a Habermasian citizen could be part of a Rawlsian overlapping consensus. While Rawls does raise the possibility of citizens who are committed to political values in themselves, rather than because they find a basis of justification in their comprehensive doctrines, this becomes less prominent in Rawls's approach and seems inconsistent with it. Rawls does not seek to accommodate the idea that principles of justice could be validated in the intersubjective domain of moral discourse, as opposed to finding a justification in their congruence with a pre-political ethical standpoint. Rawls relies on a transcendent motivational basis, comprehensive doctrines claiming a transcendence that is sublimated in the acceptance of a political conception in an overlapping consensus. Citizens recognise their reasonable transcendent conceptions as having an immanent justification in accordance with a sense of justice that is common amongst what formerly appeared to be incompatible comprehensive doctrines, and which is secured by the background of a shared social world.

If Rawls's social ontology is communitarian, Rawls's conception of philosophy is metaphysical, in Habermas's sense, in its aspiration to provide guidance for a form of life as a whole. Here too there is a parallel with Hegel, since

something like Hegel's position is implicit in the method of reflective equilibrium . . . Expressed in Hegelian language, the intended effect of Rawls's philosophy is both affirmation of 'actual' political institutions and a recognition of how existing institutions can be brought closer to actuality (Neuhouser 2000: 328-29 n.39).

Actuality here refers to 'existing social reality as reconstructed within rational (philosophical) thought'. Ideas, and ideals, already found, albeit in imperfect form, in existing institutions are actual in the sense that they are 'not externally imposed upon, but already belong to, the existing object of criticism' (Neuhouser 2000: 258). The process of reflective equilibrium 'seeks to discover a shared rational basis for political community – indeed, an incipient if latent sense of community waiting to be brought to light, elaborated, and self-consciously strengthened through reflection' (Doppelt 1990: 49).

In an overlapping consensus, a ‘political participant’s desire to act on publicly justifiable grounds is refracted through the political observer’s recognition of the fact of reasonable pluralism and emerges as a desire to avoid ideological controversy on fundamental matters, that is, to avoid being “unreasonable” (McCarthy 1994: 60). What is at issue is the authority of the third perspective of the original position that this refraction goes through. Until the achievement of a reasonable overlapping consensus, citizens lack a perspective from which to consider the best available form of social unity that they are able to achieve. The original position is used to formulate a conjecture about the most reasonable principles of social unity available to ‘us’, and the vindication of that conjecture through an overlapping consensus reflexively justifies its use as in endorsing a political conception citizens endorse the terms of the original position through which it is justified. Citizens can accept or reject the terms of the original position but not the original position itself as a philosophical device of representation.

As a ‘philosopher-participant’, Rawls experiences the fact of reasonable pluralism and, in the guise of a self-reflective moral theorist, as a ‘philosopher-observer’ constructs a political conception of justice that through reconciling us to our social world would allow us to affirm that this is indeed *reasonable* pluralism. But citizens as participants are denied such an immanent progression from the experience of reasonable pluralism to a philosophical awareness of its reasonable basis. The burdens of judgement are experienced purely as an epistemological limitation before such time as citizens endorse a political conception and see them as constitutive of an ideal of citizenship. It is only from the perspective of the ‘philosopher-observer’ that there is an awareness of the sense in which the search for the most reasonable basis of social unity continues indefinitely.

McCarthy identifies the right problem, but misinterprets Rawls, when he says:

Philosophical accounts of reason have to pass what might be called a ‘reflexivity test’: they must be applicable to themselves without producing self-referential contradictions. On this score, I think there is a tension in *Political Liberalism* between the postfoundationalist line of argument meant to convince us, the readers, of a certain conception of public reason and the thesis that reasonable citizens may embrace it for whatever religious, metaphysical, or other comprehensive reasons they find acceptable (McCarthy 1994: 61 n.29).

McCarthy wrongly assumes that Rawls aspires like Habermas to a postconventional conception of morality in Kohlberg’s sense. Kohlberg uses Rawls’s original position to exemplify the highest developmental stage of autonomous, or principled, moral judgement (Stage 6), the veil of ignorance representing ‘a statement of the fundamental

formal condition(s) of the moral point of view, the conditions of impartiality or universalizability' (Kohlberg 1981: 197). Habermas goes further, arguing that the autonomous stage of universal moral principles calls into question the existence of moral philosophy as traditionally conceived. Here Habermas follows McCarthy, and we can take McCarthy's remark regarding Rawls to derive from his observation that

Kohlberg's account places the higher-stage moral subject, at least in point of competence, at the same reflective or discursive level as the moral psychologist. The subject's thought is now marked by the decentration, differentiation and reflexivity which are the conditions of entrance into the moral theorist's sphere of argumentation. Thus the asymmetry between the prereflective and the reflective, between theories-in-action and explication, which underlies the model of reconstruction begins to break down. The subject is now in a position to argue with the theorist about questions of morality (McCarthy 1982: 74, quoted in MCCA: 173).¹¹¹

At this final stage, the asymmetry between the scientific observer and the participant is overcome. The object of analysis becomes a subject in discourse. The reconstructive scientist stands within the open horizon of a research process whose results he cannot foresee. The moral reasoning of postconventional moral subjects 'no longer mirrors, that is, prereflexively expresses, a pretheoretical knowledge but rather explicates potentially theoretical knowledge' (MCCA: 175).

Rawls, on the other hand, does not understand his readers as possessing of a postfoundationalist consciousness that allows entrance into the moral theorist's sphere of argumentation, or so I will argue. Indeed, Rawls adopts a perspective analogous to a psychoanalytic therapist, conjecturing that

'underneath' the citizens' consciously held views of justice is . . . a deeper, more or less latent, commitment to and belief in the 'fundamental' idea of justice as fair cooperation amongst free and equal persons . . . The overlapping consensus sought by Rawls is understood as one which occurs *as a result of* this reconstructive, therapeutic activity (Neal 1997: 109-10).

I address this issue further in the next section, and we will need to see whether Rawls can avoid the problems that, as we saw in Chapter 2, led Habermas to reject an analogy between critical theory and Freudian psychoanalysis and make a paradigm shift from the philosophy of the subject to the philosophy of language.

¹¹¹ See also Gadamer (1976: 42).

7.3 The *Deus ex Machina*

Paul Ricoeur frames these issues in their essential terms. Ricoeur correctly identifies the circularity of Rawls's approach, which begins by setting out the two principles of justice as fairness and then argues that they would be chosen in the original position. The argument is a progressive systematization, or rationalization, of our preunderstandings about justice. In reflective equilibrium we know the premises from which our judgements of justice are derived. 'Rawls's whole book can be thus considered as the search for this reflective equilibrium' (Ricoeur 2000a: 55). In asking whether a procedural theory of justice is possible, Ricoeur's thesis is that 'a procedural conception of justice at best provides a rationalization of a sense of justice that is always presupposed' (2000a: 50). Ricoeur concludes that this shows the difficulties of self-foundation presupposed by seeking to liberate the deontological viewpoint of morality from the teleological perspective of ethics (Ricoeur 1995: 238). However, while such a self-grounding procedural conception is something to which Habermas aspires, Ricoeur goes wrong in suggesting Rawls shares this aspiration. I take up whether Habermas can carry it through successfully in Chapter 10. For now, though, I take up Ricoeur's suggestion that the way that 'Rawls's extraordinary construction borrows its underlying dynamics from the very [sense of justice] it claims to engender by its purely contractual procedure' (Ricoeur 2000a: 38) aims at the earthly solution to Rousseau's paradox of the legislator.

What has been called the 'political circle' is a problem of utopianism that presents itself in its most acute form as a problem of founding. It presents itself for Plato, and for Cohen within a Platonic paradigm. A socialist ideal of justice is the 'no-place' of Plato's *Republic*, or the intrinsic value of Moore's Platonic atomism, that while it can never be realized can nevertheless provide a personal ideal. In our nonideal world, the role of law, like in Plato's *Laws*, is to provide a second-best, with rules of regulation instrumental to the achievement of as much justice as is possible in the circumstances. In the republican tradition, this problem, addressed by Abbe Sieyès and Rousseau, is one of the authority for founding a republic. It is 'a chicken-and-egg circle that presses us to begin the work of democratic politics in medias res' (Honig 2007: 2). Rousseau makes recourse to the figure of the Lawgiver or Legislator in response to the problem of utopianism and the political circle it generates:

For a nascent people to be capable of appreciating sound maxims of politics and of following the fundamental rules of reason of State, the effect would have to become the cause, the social spirit which is to be the work of the

institution would have to preside over the institution itself, and men would have to be prior to laws what they ought to become by means of them. Thus since the Lawgiver can use neither force nor reasoning, he must of necessity have recourse to an authority of a different order, which might be able to rally without violence and to persuade without convincing (Rousseau 1997 [1762]: 71).

We can work back to this problem, and its relevance to Rawls, from a more intuitive problem. The problem of the political circle is also a problem of re-founding, or of the ongoing authority of constitutional principles to govern subsequent generations. Rawls rightly compares Habermas's concerns about 'reigniting the radical democratic embers in a just society' with Thomas Jefferson's concern about the right of dead constitutional founders to bind those now living. This, as Hannah Arendt put it, is 'the seemingly insolvable perplexity of a revolutionary spirit that strives to establish a constitutional government. The perplexity is how to house a revolutionary spirit within a permanent regime' (PL: 408 n.45). Rawls's curt riposte is that to feel injustice that one is not able, like the founders, 'to begin the world over again' is 'entirely misplaced and cannot sensibly be entertained. (I might as well spend my life whining that I am not Kant, Shakespeare, or Mozart.)' (PL: 408 n. 45). Rawls accepts the basic model of Bruce Ackerman's dualist conception of constitutional democracy.¹¹² On this view, the constituent power of the people sets up a framework, the constitution, to regulate the ordinary power of day-to-day politics.

Higher law is the expression of the people's constituent power and has the higher authority of the will of We the People, whereas ordinary legislation has the authority, and is the expression of, the ordinary power of Congress and of the electorate (PL: 231).

While the Supreme Court in protecting the higher law is the institutional exemplar of public reason, Rawls argues that this does not place the court 'above' the people: 'The constitution is not what the Court says it is. Rather, it is what the people acting constitutionally through the other branches eventually allow the Court to say it is' (PL: 237).

This allows us to 'adjust basic constitutional values to changing political and social circumstances' (PL: 238). Here Rawls can argue against Habermas that constitutional protection of basic rights is not prior to will formation since 'the liberties of the moderns are subject to the constituent will of the people' (PL: 406). In denying that persons' autonomy is thereby restricted, Rawls asks rhetorically:

¹¹² But on Rawls's differences with Ackerman, see PL: 238-39; Freeman (1992).

Are the citizens of Rousseau's society of *The Social Contract* never fully autonomous because the Legislator originally gave them their just constitution under which they have grown up? . . . How could the Legislator's wisdom deprive citizens of the insights they have assimilated for themselves over generations? Why can those insights not be assimilated by citizens from their reflections and experience with institutions and as they come to understand the grounds of the constitution's design? (PL: 402).

There are many grounds on which from a Habermasian perspective one might wish to question Rawls's substantivist republican conception of the will of the people. For the moment, though, I restrict my attention to engaging Rawls's position on its own terms and focus on Rawls's unsuccessful attempt to sidestep the Rousseauian problem of the political circle.

It is not surprising that *TJ* should have been greeted with the thought that 'Once again a "Legislator" has appeared in our midst' (Chapman 1975: 593). Part of its enduring significance is that it promised a re-founding as a result of having captured the *Zeitgeist*, 'the new moral sensitivity, that marked the post-World War II period', the era of the civil rights movement and the Vietnam war (Audard 2007: 2). But Rawls never faces head on the problem of the standpoint which he himself occupies, and the question of by what authority he offers his political conception of justice. This is not to agree with Habermas that Rawls claims the authority of an expert, but more fundamentally to suggest that the structure of Rawls's approach makes recognition of this question of authority inevitable but answering it impossible.

The idea of an overlapping consensus, as we have seen, seeks to explain how a well-ordered society might come about. The problem faced in the formation of an overlapping consensus is similar to that faced by Rousseau in understanding the idea of a social compact. It is a problem of finding a form of association in which social unity is achieved, but in which a person 'obeys no one but himself, and remains as free as before'. This transformation in the way that we understand our freedom exchanges natural freedom for civic and moral freedom. Similarly, in the formation of an overlapping consensus, our rational pursuit of our conception of the good, given by our comprehensive doctrine, is transformed when we employ our sense of justice in recognizing the reasonable conditions, given by a political conception of justice, which establish a social space in which all persons can pursue their conceptions of the good on terms reciprocally justifiable to each person. The social compact achieves social union; it makes possible society as a social orchestra. Justice as fairness 'defines the conditions under which the spontaneous coherence of the aims and wants of individuals is neither coerced nor contrived but expresses a proper harmony consistent with the ideal good' (TJ: 281/249).

But how is political philosophy to bring this about consistent with respecting our autonomy? For:

Political philosophy cannot coerce our considered convictions any more than the principles of logic can. If we feel coerced, it may be because, when we reflect on the matter at hand, values, principles, and standards are so formulated and arranged that they are freely recognized as ones we do, or should, accept. We can use the original position to further this recognition. Our feeling coerced is perhaps our being surprised at the consequences of those principles and standards, at the implications of our free recognition (PL: 45).

The original position gives content to public reason in such a way that when principles of justice are endorsed in accordance with public reason they are endorsed autonomously. Rousseau asks:

How can it be that they obey and no one commands, that they serve yet have no master; all the freer in fact than in apparent subjection, no one loses any more of his own freedom than might harm someone else's?

His answer is that

These marvels are the work of law. It is to law alone that men owe justice and freedom. It is this salutary organ of the will of all that restores in [the realm of] right the natural equality among men. It is this *celestial voice that dictates the precepts of public reason to every citizen, and teaches him to act in conformity with the maxims of his own judgement, and not to be in contradiction with himself*. It alone is also what the chiefs should cause to speak when they command (Rousseau 1997 [1755]: 10, emphasis added).

Rawls describes Rousseau's figure of the lawgiver, whom Rousseau describes as 'the mechanic who invents the machine', as a fictional figure, a *deus ex machina*, required to answer questions of moral learning and stability, but also of the historical origins of the social compact (LHPP: 237-41). We may note here Rawls's suggestion that Rousseau's society of the social compact could come about in a similar way as Rawls envisages the society of justice as fairness coming about, through the way in which after the Wars of Religion religious toleration was accepted as a modus vivendi, providing the basis for a learning process that can lead to an overlapping consensus that generates its own support (LHPP: 240). Rawls's resolution of the problem of stability involves the perspective of a similar *deus ex machina*. It is the perspective required to provide the philosophical assurance that a political conception expresses the most reasonable basis of social unity available to us. And although from such a perspective we may only *hope* that reasonable comprehensive doctrines can support such a

conception, this is the same perspective that is required to identify reasonable comprehensive doctrines in the first place and which justifies regarding those that cannot be part of an overlapping consensus as unreasonable.

It is true, as Rawls emphasises, that we must not confuse the artificial parties who inhabit the device of representation of the original position with citizens in political society. But Rawls does seem to equate the standpoint of his readers with that of citizens committed to comprehensive doctrines and situated within the background culture of civil society. Rawls subscribes to a democratic view of political philosophy according to which it is 'part of the general background culture of a democratic society', addressing itself to citizens. However, he envisages that 'in a few cases certain classic texts become part of the public political culture', part of society's fund of basic political ideas that exist alongside the constitution (LHPP: 3). The audience for Rawls's work corresponds to its two Rousseauian stages: the legislative task of the *Social Contract* and the educative task of *Emile*, politics and pedagogy. Just as Rousseau was educating Emile to be a citizen among those who are to be (Cassirer 1954: 123), so Rawls's work seeks to educate citizens into an understanding of themselves that would seem to require institutional reform, and the acceptance of justice as fairness in the public political culture, to come to fruition. However, when political philosophy plays a role in constitutional politics, does it not risk operating above the heads of citizens? How can such constitutional issues be justified in terms of public reason if their content is based on a political conception that defines an ideal of public reason and they are determined by an institution, the Supreme Court, which is the exemplar of public reason? Similarly, when political philosophy has an 'educative role as part of the background culture' and contributes to how citizens 'learn from civic society its fundamental conceptions and ideals before they come to democratic politics' (LHPP: 5-6), does it not risk operating behind the backs of citizens? In Rawls's social ontology there is no social space in which the conception of justice as fairness could be read as a philosophical doctrine and assessed by citizens on these terms. Rawls identifies no basis upon which philosophy can transcend itself.

Neal detects a tension between 'heroic' and 'benign' interpretations of the role of the theorist and considers the following possibility:

The textual elements supporting the benign reading are the sheep's clothing within which are cloaked the ambitions of the heroic founder, who cannot present himself as such precisely because he knows there can be no authority for what he is doing (attempting to do) other than the success which will create it retrospectively. As with Rousseau's Legislator, we find here '...simultaneously two things that seem incompatible: an enterprise

threatening to surpass human powers, and, to execute it, an authority that is a mere nothing.' But can *that* be Rawls? (Neal 1997: 127, quoting Rousseau).

We are faced again with the perplexities of Rawls's 'audacious humility'. This is well captured in an excerpt from a 1986 conversation between Rawls and Habermas (quoted in Lloyd (1994)):

Habermas to Rawls:	But surely you, as a philosopher, must claim truth for your theory.
Rawls to Habermas:	I think that we, as students of philosophy, should be allowed to claim for our theories whatever we think most appropriate.

Little rests on the veracity of this recollection since it is a distillation of sentiments readily identifiable in the later debate between Rawls and Habermas. Rawls's humility is apparent in his disavowal of any claim to being a philosopher. Rawls's self-image is as a *student* of philosophy, concerned like citizens generally with making sense of the terms of our collective life. But there is audacity in the reconstructive attitude that this licenses in interpreting the philosophical canon in a manner suitable 'for us', and in the practical constructive role consequently allotted to philosophy.

This all goes some way towards meeting Habermas's criticism that Rawls presupposes a standpoint of philosophical expertise. But fundamental difficulties remain, and Rawls's response to Habermas's critique is revealing:

In justice as fairness there are no philosophical experts. Heaven forbid! But citizens *must*, after all, have some ideas of right and justice in their thought and some basis for their reasoning. And students of philosophy take part in formulating these ideas but always as citizens among other (PL: 427, emphasis added).

How are we to understand this 'must'? We can see Rawls as 'engaged in a sophisticated and rhetorically nuanced effort to foster the conditions of agreement needed to carry the convictions his arguments are supposed to support' (Scherer 2006: 358). We see this when Rawls says:

It is important to distinguish between three points of view: that of the parties in the original position, that of citizens in a well-ordered society, and finally, that of ourselves – of you and me who are elaborating justice as fairness and examining it as a political conception of justice (PL: 28, see also 293).

Rather than clarifying the points of view involved, this seems intentionally to blur them. In what sense is a reader of *PL elaborating*, alongside Rawls, justice as fairness? In what sense is Rawls in the same position of a reader of *PL examining* a political conception of justice? Rawls's standpoint of philosophical authority can only find its justification in fostering political conditions that allow the endorsement of the political conception derived from it, in Rawls showing *his* understanding of justice to be *ours*.

Finally, we could push this idea of the 'God in the machine' even further, for in Rawls's mature work justice as fairness could even be said to play a role analogous to God in Rawls's senior thesis, the achievement of reflective equilibrium in an overlapping consensus being strikingly similar to Rawls's early description of religious conversion. Just as there is no formation of an overlapping consensus without individuals endorsing a political conception of justice, so, because God restores man to community, man must wait for God to speak to him before he can communicate with the 'other' in community (MSF: 224-25). Conversion has a dialectical quality, 'a being burst in upon and being bent back':

the man who is elected and restored to sonship before God is not thereby lifted out of the earthly community; he is not abstracted from his fellow men and assumed into heaven. Rather, he is planted more firmly into that earthly community (MSF: 233, 244).

Endorsing a political conception of justice in an overlapping consensus is ultimately seen to rest on a leap of democratic faith. When it comes down to it, Rawls seeks to engage his readers' hearts, the desire for purity of heart that 'if one could attain it, would be to see clearly and to act with grace and self-command' (TJ: 587/514). While this might be thought an attractive aspect of Rawls's approach (Weithman 1994, 2009), we should explore whether a more adequate reconciliation of realism and a utopian impulse is possible.

7.4 Post-Rawlsianism

In starting out from Rousseau's realistically utopian idea of 'taking men as they are and laws as they might be', Rawls never adequately reconciles Kantian and Hegelian ideas, remaining caught between forward-looking transcendent faith and backward looking concern with immanent potential. Before going on to argue that Rousseau's theodicy has a role analogous to Kant's postulates of practical reason in its concern with what we may hope for, Neuhouser claims:

Even if [Rousseau's] account of the redemptive potential of *amour-propre* cannot reconcile us to present reality – cannot guarantee that the promise of good that is hidden in the evils of our actual circumstances is or ever will be realized – there may still be a kind of reconciliation it offers: affirmation of *the world in its basic structure* (Neuhouser 2008: 6).¹¹³

This is the sense in which we should understand Rawls's conclusion that 'While realization is, of course, not unimportant, I believe that the very possibility of such a social order can itself reconcile us to the social world' (LP: 128). There is a suggestion of an opportunity missed, of a substantive ideal of social arrangements that is no longer possible for us, but whose very possibility means 'we can reasonably hope that we or others will someday, somewhere, achieve it' (LP: 128).

But Rawls also says that the possibility of a well-ordered society 'is not a mere logical possibility, but one that connects with deep tendencies and inclinations of the social world' (LP: 128). Catherine Audard suggests that 'Philosophy of history is possibly the missing link which would demonstrate the full potential of Rawls' realist[ic] utopia' (Audard 2007: 78). Ultimately Rawls does not offer a critical social theory and succumbs to an impotent idealism. For Habermas, on the other hand, a reconstruction of the idealizing presuppositions of communicative action reveals 'a glimmer of symmetrical relations marked by free, reciprocal recognition. But this idea must not be filled in as the totality of a reconciled form of life and projected into the future as a utopia' (PT: 145). A proceduralist concept of rationality 'cannot sustain utopian projects for concrete forms of life as a whole' (RR: 87).

Political philosophy is post-Rawlsian in the sense that, as Nagel put it, Rawls 'changed the subject', but it is increasingly argued that political philosophy should move beyond Rawls. In advocating a turn to a Habermasian framework, I will seek to show that this approach more effectively takes forward Rawls's concerns than other approaches that have been dubbed post-Rawlsian. Post-Rawlsians, who include David Peritz, Joshua Cohen and Anthony Laden, seek to 'extract the core requirement of reasonable acceptability from Rawls's account of public reason, which is inextricably bound up with his account of political liberalism' (Peritz 2004: 268).¹¹⁴ A shared assumption is that while neither a political conception of justice nor a family of political conceptions can be reasonably accepted by free and equal persons in a deeply diverse democracy, the idea of reasonable acceptability as a basis of democratic legitimacy can

¹¹³ I take the Rawlsian allusion to be deliberate.

¹¹⁴ Peritz (2004) credits this term to Charles Sabel, who used it to describe Peritz and Joshua Cohen. Peritz adds Laden to the list, and Laden notes the similarities between his approach and Bohman's.

be extracted from this context. Principles of deliberative democracy rather than principles of justice become the focus of attention.

Here I shall briefly take up Laden's position, leaving Joshua Cohen's position to be analysed in the next chapter. Laden's deliberative liberalism seeks to carry forward Rawls's conception of philosophy as defence. It seeks to 'defend our reasonable faith in the possibility of a legitimate constitutional democratic regime in the face of deep diversity' (Laden 2001: 7). In focusing on this, and Rawls's appropriation of Rousseau and Hegel, I have followed Laden's lead, but in turning to Habermas I will follow a different interpretative direction. Laden distinguishes between a natural law strand and a deliberative strand in Rousseau's theory. According to the natural law strand, the social contract that provides the structure for the general will is justified because it embodies the fundamental interests of citizens. According to the deliberative strand, on the other hand, principles are legitimate when they are supported by the general will, and the social contract is justified by its establishing the social conditions that enable people to form and be governed by the general will. Laden argues that conditions of deep diversity undermine the natural law strand but provide support for the deliberative strand (2001: 34-35). He presents Hegel as a critic of the natural law strand and a supporter of the deliberative strand. However, with Hegel too Laden differentiates between immanent and transcendent elements, separating out a social freedom argument found in the *Philosophy of Right* that he accepts from a metaphysical freedom argument that he rejects (2001: 49).

In using this deliberative strand as the basis for an account of democratic deliberation, Laden recognises that his own understanding of public reason departs from Rawls's 'official' definition of public reason where its content is given by a political conception of justice, adopting instead a general definition of public reasons, or 'we'-reasons, as those we can reasonably expect other citizens to endorse (2001: 116). "We"-reasons are reasons deliberators presume to be authoritative for those with whom they comprise the plural subject engaged in deliberation' (2001: 96). However, here fostering deliberation risks becoming an end in itself, severed from a conception of individual and collective autonomy. In contrast to Laden, we could see the fact that there are grounds for faith in the stability of democratic societies under conditions of diversity as evidence that the corollary of increased pluralism is an increased scope for, if not the full realization of, the maintenance of social order through discursive practices that appeal to context-transcending standards of reason. This, at least, is the direction I will take in

seeking a more adequate utopian realist defence of reasonable faith in the competence of human reason.¹¹⁵

7.5 On the Very Idea of a Comprehensive Doctrine

Looked at from one perspective, the paradox of Rousseau's social pact is that it has to bring about the conditions that will make itself possible. But from the opposite perspective, 'what makes . . . transformation possible at the same time makes it unnecessary' (Lassman 2003: 52). Is it really the case that citizens require ideas of right and justice formulated by philosophers, amongst others, since otherwise they will be confined to the perspective of their comprehensive doctrines? Habermas asks 'could the religious conflicts [of the Wars of Religion] have been brought to an end if the principle of tolerance and freedom of belief and conscience had not been able to appeal, with good reasons, to a moral validity independent of religion and metaphysics?' (HR 1: 67). Is it actually that philosophy, in applying the principle of toleration to itself, is catching up with social transformations rather than seeking to bring them about?

Citizens as Rawls conceives them within the terms of justice as fairness do not take on real flesh and blood form, remaining identified with the comprehensive doctrines they hold. But are there really utilitarians and Kantians for whom these doctrines play a role in their lives analogous to the role of religious faith for a religious believer? In translating his approach to reconciling competing methods of ethics in moral theory into political terms, Rawls's overlapping consensus retains the character of a reconciliation of theories. The position that citizens are able to adopt is presented as a choice between 'unreasonable insistence on the truth of one's beliefs and reasonable avoidance of claims to truth', 'sectarianism and civility' (McCarthy 1994: 62).

However, whilst a comprehensive doctrine may not be able to adopt a reflective attitude towards itself, there is no reason to believe that a citizen committed to a comprehensive doctrine is not able to do so. We can take up the question of whether one can make sense of the point of view that Rawls presupposes for deriving substantive principles capable of achieving reconciliation through considering Donald Davidson's (2001) critique of Quine's scheme-content dualism as the third dogma of empiricism. Rawls's understanding of how comprehensive doctrines provide the frameworks through which citizens make sense of their social world is relevantly analogous to Quine's idea of conceptual schemes. Just as we may ask about the tenability of the very idea of

¹¹⁵ Similar considerations apply to Bohman's pragmatic critical theory (Bohman 1996), the challenge of pluralism having also led to a pragmatic turn in critical theory (Rehg and Bohman 2001).

differing conceptual schemes interpreting a common world, so we can ask about the tenability of the idea of comprehensive doctrines that may be reconciled on the basis of a political conception of justice. The following insight will guide my interpretation of Habermas in relation to Rawls:

[T]he move from Fichte to Hegel is analogous to the move from Quine to Davidson: Quine starts with an isolated subject, naturalistically construed, who is then stimulated by his environment to respond by ‘positing’ certain items to account for them; Davidson, unlike Quine, begins with a ‘triangulated’ conception of subjects in intersubjective communication about the world (still naturalistically construed), interpreting each other fundamentally in terms of their common links to objects in that world (Pinkard 1999: 230 n.3).

As we will see in Chapter 9, Habermas begins from such a Davidsonian model, and the challenge he seeks to overcome is one of showing how consistent with this naturalistic picture, a context-transcending conception of communicative practical reason may still be defended. The limitation of Rawls’s metaphor of social unity as a social orchestra is that symphony orchestras have conductors that mediate the formation of intersubjectivity. But in rejecting this metaphor, we need not adopt Cohen’s jazz band. Alongside his orchestra analogy, Rawls also appeals to the way in which a competitive game relies on the players sharing a ‘public desire to execute a good and fair play of the game’ (TJ: 525-26/461). Gadamer likens the formation of intersubjectivity in language games to the activity of play (Gadamer 2003 [1960] : 101-110), and this provides a starting point for understanding Habermas’s approach.

In this reconfigured economy of philosophical ambition, the corollary of the more developed account of the interpretative accomplishments of moral subjects is a more modest conception of the possible role of substantive rules of social cooperation. The need for this move is well brought out by Stanley Cavell.¹¹⁶ Cavell draws attention to the significance of Rawls’s idea that the institutional realization of a conception of justice means that persons can say to one another in their everyday lives that their actions are above reproach (Cavell 2005: 171; 1990: 18, 113). In contrast to Cohen’s structurally similar interpretation, Cavell correctly identifies the connection between Rawls’s focus on rules of social cooperation to regulate the basic structure of society and his aspiration for reconciliation to our social world. However, as well as diagnosing a problematic aspect of Rawls’s architectonic, Cavell can also help us to locate its source. Cavell’s interpretation of Rawls builds upon an earlier critique of Rawls’s understanding of rules in ‘Two Concepts of Rules’ (Cavell 1999: 292-313). It is now apparent that Rawls’s

¹¹⁶ On Cavell’s critique of Rawls, see Mulhall (1994, 1997).

dichotomous strategy of reconciling the competing pull of forward-looking and backward-looking concerns by making them apply to two different situations, those of the legislator and the judge, persists throughout his work. The philosopher adopts the perspective of the legislator, recognising a practice conception of rules, and citizens from the perspective of their comprehensive doctrines adopt the perspective of the judge, applying a summary conception of rules. However, following Cavell, and his appeal to Wittgenstein and J.L. Austin (1975), we can say that to understand the true rationale for rules requires adopting a performative attitude through which participants understand a rule in the activity of applying and reproducing it. Habermas rejects the view that political obligation derives from seeing democracy as a current counterpart of a promise of the Founding Fathers. Promises are illocutionary acts that establish interpersonal relations and it is a consequence of persons' autonomy that the duties they generate can only have a specific content and not the binding character of obligation as such (BFN: 548 n. 76). In Part V, I will trace the connections in Habermas's approach that lead from the illocutionary force of performative utterances in reflexive social practices to a political conception of the constitution as a project.

Conclusion

Nozick greeted the arrival of *TJ* with the observation that 'Political philosophers now must either work within Rawls' theory or explain why not' (Nozick 1974: 183). But Nozick's approach to explaining why he would not, like Cohen's, testifies to a persistent failure to address Rawls's methodology in its own terms (Delaney 1977). Critics of Rawls have tended to exist in a symbiotic relationship with his work, taking it to legitimate a field of enquiry at the same time as rejecting it. But as Rawls says, 'Objections by way of counterexamples are to be made with care' since the real question is which proposed approach 'is the best approximation overall' (TJ: 52/45).

The problem is, though, that with political liberalism Rawls aspires to 'the political philosophy to end all political philosophies' (Kukathas 1990: 151). Rawls's conception of justice cannot 'see itself as just one among many competing political philosophies. Its aim, ultimately, is not to challenge or repudiate such competitors but to subsume them, by formulating a standpoint which will command their allegiance' (Kukathas 1990: 149). However, particularly now that our perspective on Rawls's work has become as much about its place in the history of political thought as about how it defines the terms of contemporary political theory, there is another possibility: a hermeneutic strategy of appropriation which Rawls would be the first to recognise. To

critique Rawls we need a standpoint beyond Rawls's theory, but such a standpoint can only emerge from a reconstructive approach to Rawls's project that identifies its limits and seeks to go beyond them.

There is some pathos in the way that the cunning of reason has meant that Rawls's work has contributed to trends of specialization and rationalization in philosophy that serve to undermine the ability of justice as fairness to perform the role that it sets for itself. As Habermas argues, philosophy should not regress 'into the self-sustaining discourses of an academic discipline', but neither can it now imbue our lives with meaning; it cannot 'slake the thirst of the sons and daughters of modernity by providing some surrogate for the lost certainties of religious faith' (TaJ: 283, 289). Philosophy situates itself within this world at the same time as it interprets it, finding its place as a specialized science amongst others without thereby fully renouncing its role as the public guardian of democratic rationality. In recognising the unrealizability of reconciliation to our social world, philosophy abandons the effort to sustain reasonable faith in a substantive ideal of a well-ordered constitutional democratic society in favour of preserving rational faith in the capacity to maintain and develop a just social order through procedures of communicative reason.

PART IV. BEYOND ‘RAWLBERMAS’

8. The Democracy/Contractualism Analogy Beyond ‘Rawlbermas’

The publication of Rawls’s *PL* in 1993 coincided with the ‘deliberative turn’ in political theory (Dryzek 2002) as part of which there has been a shift away from theories of justice and towards theories of political legitimacy (Peter 2007: 1; Estlund 1996). The overall aim of this chapter is to consider the methodological differences between G.A. Cohen, Rawls and Habermas in the context of debates about the relationship between justice and deliberative democracy. To this end I will focus on the approach of Estlund, which, I will argue, applies a Cohenite Platonic paradigm of justice to questions in democratic theory. Estlund has engaged with the work of both Rawls and Habermas, arguing that those interested in either Rawls or Habermas ‘ought really to be interested in both’ (Estlund 1992: 694). But Estlund criticises the deliberative turn’s shift away from a Rawlsian concern with justice to a Habermasian concern with legitimacy. More specifically, he criticises the tendency of theorists to rely upon a democracy/contractualism analogy in which democratic institutions are thought of as ideally mirroring a contractualist procedure for choosing principles of justice. The problem, as Estlund sees it, is that the democracy/contractualism analogy consists in a ‘flight from substance’, a denial of procedure-independent standards that has been the characteristic ambition of recent normative democratic theory (Estlund 2008: 97). In beginning to clarify the differing understandings of proceduralism in Estlund, Rawls and Habermas, this chapter sets the stage for my interpretation of Habermas in the next chapter and my discussion of the issues of procedure versus substance in the Rawls-Habermas debate in the final chapter.

As part of considering Estlund’s argument against the democracy/contractualism analogy, it will also be necessary to assess the sense in which Rawls and Rawlsian theorists propound a conception of democratic deliberation. In looking at the methodological debate between G.A. Cohen, Rawls and Habermas in the context of deliberative democracy, this chapter is a bridge between my discussion of Rawls and my discussion of Habermas. For although with his idea of public reason Rawls is commonly thought to have joined the deliberative turn (Bohman 1997) – stating explicitly in ‘The Idea of Public Reason Revisited’ that he understood a well-ordered constitutional democracy to be a deliberative democracy (CP: 579) – it is Habermas who is the ‘father of deliberative democracy’ (Chambers 2002: 169).¹¹⁷ Alongside Habermas himself, the

¹¹⁷ See also Elster (1998: 1). Habermas’s influence is evident in the canonical early statements of deliberative democracy by Elster (1986); Manin (1987); J. Cohen (1997b).

proponents of the democracy/contractualism analogy who Estlund discusses have sought to integrate a Habermasian conception of discourse into a Rawlsian architectonic. The most important intermediary between Rawls and Habermas has been Joshua Cohen, who has sought to connect ‘Rawls’s emphasis on substantive justice and Habermas’s emphasis on legitimate procedure’ (Estlund 2001: 5). However, Estlund takes this Habermasian influence to be significant enough to have led J. Cohen to reject normative standards prior to or beyond democratic procedures, and consequently he labels both Habermas and J. Cohen as advocates of ‘deep deliberative democracy’ (Estlund 2008: 87-93; 2005: 222-27). Other proponents of the democracy/contractualism analogy such as Barry have sought to apply Scanlon’s contractualist moral theory to issues in political and democratic theory. It has been argued that Scanlon’s contractualism is a ‘bridge between liberal theory and critical theory’ since Scanlon, unlike Rawls, endorses normative justification through conversation rather than contract (Chambers 1996: 79; see also Baynes 1992: 115-18).

I will argue that attempts to build bridges between Rawls and Habermas are beset with difficulties. Scanlon’s moral theory continues to operate within a Rawlsian framework and the bottom-up Kantian Republicanism of Habermas’s political theory of *discursive* democracy contrasts significantly with the top-down liberal constitutionalism of *deliberative* democracy.¹¹⁸ While on this basis I endorse Estlund’s rejection of J. Cohen’s democracy/contractualism analogy, it is for the opposite reason to Estlund’s view that it is objectionably procedural. J. Cohen’s view remains substantive in a Rawlsian sense and represents an untenable halfway house between Rawls and Habermas. It subscribes to the Kantian paradigm of the relationship between theory and practice, pursuing neither Rawls’s Rousseauian-substantive justice-based approach towards the problem of self-grounding nor Habermas’s Rousseauian-procedural discourse-based approach. J. Cohen has contributed to the existence in the literature of what I will refer to as the figure of ‘Rawlbermas’, a hybrid of Rawls and Habermas against which I will argue. Estlund’s approach is to argue that participants in democratic deliberation must directly address what is demanded by the truth about justice, separating a substantive Rawlsian ideal of justice from a Habermasian procedure of real discourse. This returns to the Platonic paradigm of the relationship between truth and politics that I rejected in examining G.A. Cohen’s interpretation and critique of Rawls. I will argue, by contrast, that in failing to sufficiently distinguish Habermas’s position from J. Cohen’s, Estlund mischaracterises the nature of Habermas’s proceduralism and leaves this alternative elaboration of a Rousseauian paradigm unexplored.

¹¹⁸ This distinction follows Dryzek (2000, 2002).

8.1 The Democracy/Contractualism Analogy

As Estlund presents it, the democracy/contractualism analogy is the view that democratic procedures can produce outcomes that are right by contractualist standards, and therefore ‘track’ justice, insofar as they are structurally similar to the procedure in a hypothetical contractualist situation (Estlund 2008: 239). While this has the advantage of avoiding potentially controversial claims about the nature of justice, Estlund argues that the analogy is implausible. Justice cannot be promoted through the ‘invisible hand’ of democratic deliberation, but rather depends upon citizens engaged in deliberation directly asking the ‘primary question’ of what justice demands (Estlund 2003b: 387).

To properly address Estlund’s critique it will be helpful to situate it within the context of Estlund’s overall approach to democratic theory. Estlund (1) establishes a taxonomy of procedural and substantive justifications of democracy; (2) critiques the ‘deep deliberative democracy’ of J. Cohen and Habermas for its reliance on the democracy/contractualism analogy; and (3) defends a theory of epistemic proceduralism, according to which democracy is justified not only by its procedural legitimacy or fairness, but also by its ability to track procedure-independent substantive standards of justice. The remainder of this section sets out Estlund’s philosophical framework and briefly introduces his defence of epistemic proceduralism. I return to Estlund’s own substantive position in 8.5 after examining his argument against the democracy/contractualism analogy.

Estlund (1997) locates J. Cohen and Habermas on a spectrum running from procedural to substantive approaches to democracy (Figure 4). Fair Proceduralism, as advocated by Robert Dahl and Stuart Hampshire, identifies democratic legitimacy simply with the fair procedure of majority rule. To this, Fair Deliberative Proceduralism, proposed by Bernard Manin and Frank Michelman, adds the requirement that citizens be given an equal, or at least fair, chance to contribute arguments and reasons to a discussion before voting. The final form of proceduralism, the Rational Deliberative Proceduralism of J. Cohen and Habermas, differs from Fair Deliberative Proceduralism in locating the value of the democratic procedure primarily in its recognising good reasons rather than securing fair access (Estlund 1997: 179). This position remains proceduralist because for Habermas ‘there are no standards that loom over the political process, policing its decisions’ (Estlund 2008: 88). For J. Cohen too democratic authority is free and self-determining. ‘It is not under any other authority, not even the “authority” of prior normative standards for better or worse decisions’ (Estlund 2008: 90).

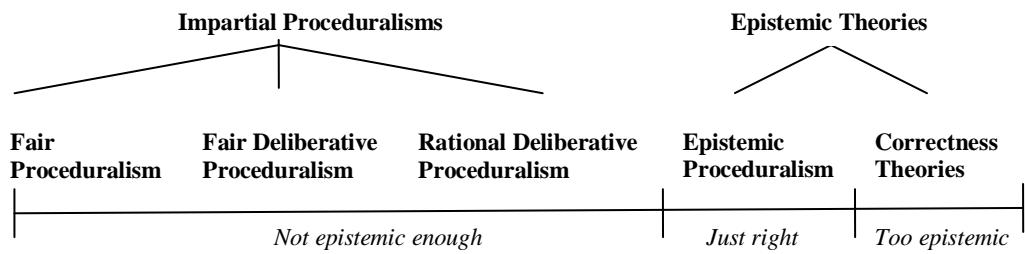


Figure 4. Procedural and Substantive Justifications of Democracy

Source: Estlund (1997: 182; 2008: 102).

But Estlund holds that even Rational Deliberative Proceduralism is untenable, since if it claims to distinguish better reasons from worse this can only be because ‘those reasons are being counted as better by procedure-independent standards’ (Estlund 1997: 179). Such standards need not be independent of *all* procedures; they could be the standards identified by an *ideal* contractualist procedure, but they are independent of the *actual* procedure of deliberation. In rejecting all ‘impartial proceduralisms’ in favour of an epistemic defence of democracy, Estlund nevertheless seeks to avoid swinging to the opposite epistemic extreme of a ‘correctness’ epistemic theory. He does follow Plato in believing both that justice is a substantive standard, at the bar of which the quality of political outcomes may be judged, and that some citizens will therefore be superior in wisdom and commitment when it comes to promoting better outcomes (Estlund 2008: 206). But Estlund (1993a) is concerned to make the idea of political truth ‘safe for democracy’ by severing the inference from claiming to know the truth to claiming an entitlement to rule. It is at this point that the procedural element of epistemic proceduralism, and the influence of Rawls, enter the picture. Taking inspiration from Rawls’s idea of the reasonable, Estlund argues that political legitimacy depends upon democratic procedures and their outcomes meeting an acceptability requirement. Epistemic proceduralism can recognise the existence of procedure-independent standards without endorsing ‘epistocracy’, or rule by those who know best. In a form of constrained optimisation familiar from rule utilitarianism, epistemic proceduralism seeks to identify the truth about justice subject to the legitimacy constraint that the democratic procedures that are used to this end must be acceptable to all qualified points of view.

8.2 Contractarianisms and Contractualisms

With Estlund's overall philosophical framework in view, as well as the general course that his own position takes in response, we can return to his critique of the democracy/contractualism analogy. The problem that one is immediately faced with is that Estlund's broad idea of a democracy/contractualism analogy – as opposed to the specific ways in which this idea is developed by Barry, J. Cohen and Habermas – is a composite construction that no one theorist has proposed in its entirety. Estlund's presentation of the analogy borrows elements from Barry and J. Cohen.¹¹⁹ It takes from Barry the idea of a 'Scanlonian original position' and from J. Cohen the idea of mirroring this initial situation in deliberative democratic institutions. In the remainder of this chapter I leave aside the broad conception of the analogy and focus on J. Cohen's specific approach. I do this because, as I will argue in the remainder of this section, it fundamentally misconstrues the nature of Scanlon's contractualism to see it, as Barry does, as providing an alternative to Rawls's contractarian use of the original position.¹²⁰ While I will ultimately reject J. Cohen's view, it is worthy of further consideration, which I turn to in the next section, as although J. Cohen draws upon Scanlon's contractualist account of motivation, his model of an ideal initial situation owes more to Habermas's ideal speech situation than to Scanlon's test of reasonable rejection.

Scanlon's contractualist formula states that 'an act is wrong if its performance under the circumstances would be disallowed by any set of principles for the general regulation of behaviour that no one could reasonably reject as a basis for informed, unforced general agreement' (Scanlon 1998: 153). Scanlon has been seen as lifting Rawls's veil of ignorance, thereby suggesting a democracy/contractualism analogy and a bridge between Rawls and Habermas. Why would it be mistaken, though, both to see Scanlon as offering an alternative to Rawls's way of justifying principles of justice and as endorsing a Habermasian commitment to discourse? Scanlon's contractualism has

¹¹⁹ Estlund (2008: 291 n.5) also identifies William Nelson as drawing a democracy/contractualism analogy, but I will not address Nelson's position.

¹²⁰ Whereas for Scanlon the contractualist idea of justifiability is constitutive of rightness (Pettit 1999), Barry (1995a) uses Scanlon's formula at an *a priori* level as a useful heuristic for focusing our intuitions about justice (Pettit 1996). At an empirical level, Barry is concerned with the extent to which social conditions approximate those of a Scanlonian original position. Here Barry endorses deliberative democracy as 'the ideal way to *implement* substantive norms of justice' (J. Cohen 2003: 136 n.37, emphasis added), rather than arguing that democratic procedures should seek to mirror a Scanlonian original position. Barry's position can therefore be reconciled with a consequentialist conception of rightness in a way that Scanlon's cannot. Barry's approach is compared to Habermas's by Rytina (1990) and Pogge (1990), but the comparison is highly strained. According to Barry (1995a: 196 n.), 'Habermas holds that we have to wait patiently until an ideal speech situation exists rather than try to anticipate what might emerge from it if it did exist'.

been interpreted in many different ways, and it will illuminate both questions to consider different interpretations with reference to the distinction between Platonic, Kantian and Rousseauian paradigms of the relationship between theory and practice introduced in Chapter 1.¹²¹

Estlund's interpretation of Scanlon, and those like Barry who draw upon him as the basis for a democracy/contractualism analogy, is based on a commitment to a Platonic paradigm. This is the same approach that is adhered to by consequentialist critics of Scanlon such as Pettit, subscribing to a division of labour between the desirable and the feasible. The problem is that from this perspective Scanlon's contractualism appears to be a redundant epicycle. Plato's 'Euthyphro question' asks whether something is pious because it is loved by the gods, or whether the gods love it because it is pious, and a similar question is asked of Scanlon's contractualism by Estlund and Pettit. The 'redundancy objection' to Scanlon's contractualism takes the form of a dilemma. Scanlon's contractors may either offer non-moral reasons as the grounds for why it is reasonable to reject a principle, but if they do then Scanlon will have failed to provide an account of moral wrongness: his contractualist machinery cannot turn non-moral reasons into moral reasons. Scanlon is after all not a neo-Hobbesian contractarian like David Gauthier, conceiving of morality as a rational self-interested bargain. However, if alternatively Scanlon's contractors offer moral reasons, then the grounds for reasonable rejection will be *antecedent* moral reasons, such as considerations of welfare or fairness, and the contractualist procedure will be otiose.

From the perspective of the Platonic paradigm, Scanlon's contractualism has been defended against the redundancy objection by Michael Ridge (2001), and it is this interpretation to which Estlund is indebted (Estlund 2008: 291 n.17). Ridge argues that while consequentialist critics assume that grounds for reasonable rejection must be agent-neutral reasons, on Scanlon's view they 'not only can be agent-relative, they *must* be' (Ridge 2001: 473). Scanlon can disarm the first horn of the dilemma by arguing that non-moral objections can have moral content, being relevant for fundamental moral judgements when they are *reasonable*. He can then disarm the second horn by arguing that agent-neutral reasons cannot be separated from agent-relative reasons: 'moral reasons "piggyback" on people's agent-relative reasons' (2001: 478). Agent relative reasons therefore both have moral content and are ineliminable. The notion of reasonable rejectability "bridges the gap" between one agent's agent-relative reason and the agent-neutral judgement that an action is wrong' (Frei 2009: 53), representing an

¹²¹ Here I broadly follow the taxonomy of Ashford (2007).

agreement made between contractors to make room for others' agent-relative reasons in coming to their all-things-considered moral judgements.

But as Tamar Frei (2009) has pointed out, Scanlon, like Rawls and Korsgaard, does not accept the distinction between agent-neutral and agent-relative reasons, and so cannot be looking to 'bridge the gap' between them. For Scanlon, all reasons arise out of a standpoint. Practical reasons are generic rather than specific, and personal rather than impersonal. Scanlon *begins* from the assumption that all rational agents are agent-neutrally valuable and intends his theory to explicate 'what is involved in recognizing and respecting the (assumed) agent neutral value of rational agency' (Frei 2009: 57). This means that Scanlon does not share with consequentialists the aim of grounding morality. His 'claim is not that we have some reason to care about the generic personal reasons of others because we, or some idealized version of ourselves, have or would agree to do so' (Frei 2009: 58). Scanlon rather seeks to provide the most phenomenologically accurate account of moral motivation. In determining amongst competing reasons which provides the strongest grounds for rejection, Scanlon does not appeal to the idea of an agreement amongst agents to settle the issue. Instead, intuitive judgement on the part of agents is unavoidable.

But if Scanlon does not seek to ground morality then an interpretation of his contractualism in terms of the Kantian paradigm will also be unconvincing. From a Kantian constructivist standpoint, O'Neill (2003b) argues that, in her terms, Scanlon's theory is constructivist rather than contractualist as it is concerned with the universalizability of reasons rather than, like Rawls's, with the fact of agreement.¹²² It is true that Scanlon follows the Kantian idea of respect for rational agency and therefore sees morality as concerning mutual recognition of the justifiability of principles between free and equal persons. But while Scanlon seeks to elucidate the normative appeal of Kant's categorical imperative in its formula of humanity and formula of the kingdom of ends variants, he rejects the motivation behind Kant's formula of universal law. Scanlon acknowledges that he departs from Kant in taking the *substantive* appeal of this ideal of justifiability as the normative foundation of morality, rather than giving it any deeper foundation, as Kant does when he links the moral law to the very idea of rational agency. Scanlon 'offers a substantive account of the normative force of morality, based on the value of a relation of mutual respect. Reasonableness is not taken to be something that can be demonstrated outside the moral point of view' (Ashford 2007).

Not for nothing does Scanlon say that the intrinsic value of justifiability to others is 'an important element in the thinking of philosophers in the contract tradition, such as

¹²² Scanlon (2003e: 424-26) rejects O'Neill's Kantian reading and acknowledges he is seeking 'Kant on the cheap' (Voorhoeve 2001: 30).

Rousseau' (Scanlon 1998: 73).¹²³ Scanlon and Rawls share a Rousseauian ideal of justifiability among free and equal persons, with Scanlon's idea of reasonable rejectability developing Rawls's idea of reasonableness as the source of moral motivation. Scanlon's domain of 'what we owe to each other', while not covering the whole of morality, is indeed broader than the scope of political principles of justice. But, *pace* O'Neill, it would be a mistake to thereby take Scanlon's contractualism to have greater critical potential than Rawls's approach. Rawls is concerned with political philosophy, a moral conception for the basic structure of society.¹²⁴ Scanlon, on the other hand, is concerned with a moral philosophy that elucidates individual morality (Scanlon 1998: 73). While Rawls is concerned with principles for the reform of the basic structure of society that meet the liberal principle of legitimacy – according to which political power is legitimate when it is exercised in accordance with ideals that all citizens can be expected to endorse (PL: 137) – Scanlon' phenomenological account assumes a background of reasonableness, analogous to Rousseau's general will, that leads agents to seek principles no one could reasonably reject. Scanlon's conception of moral theory is more modest than Rawls's (Scanlon 1992).

Scanlon's contractualism identifies a particular question as relevant to differentiating right and wrong. It is not Hobbes's question: 'Is there an alternative action that would lead to a better outcome?' And it is not Kant's question: 'Is the maxim that I am contemplating acting on universalizable?' It is rather: 'Are the morally relevant generic personal reasons for the principle that forbids the action weightier than ones that recommend rejecting the principle?' (Frei 2009). This is a question of judgement, of putting oneself in other people's shoes. While Scanlon's contractualism therefore shares a number of features with Habermas's discourse theory, Scanlon disagrees with Habermas on two important points. First, there is the question of moral motivation. As Scanlon explains, accounts of the importance of morality and its reason-giving force can be compared according to whether they are formal or substantive. Kant provides the leading example of a formal strategy, and Habermas follows Kant's lead in arguing that 'valid moral principles can be derived in argument following rules that must be presupposed by anyone who undertakes to engage in argument at all'. Scanlon rejects the strategy of appealing to considerations that are independent of particular ends, seeking instead to 'explain the reason-giving force of moral judgements by

¹²³ 'Contractualism has its roots in Rousseau, rather than Hobbes: the general will is what we would jointly will if we adopted the perspective of free and equal citizens' (Ashford 2007). Darwall (2002) obscures this distinction by contrasting Hobbesian contractarianism with Kantian contractualism, as does Southwood (2008).

¹²⁴ For this reason I follow Rawls's practice of distinguishing his own contractarianism from Scanlon's contractualism (Freeman 2007c: 36).

characterizing more fully, in *substantive terms*, the particular form of value that we respond to in acting rightly and violate by doing what is wrong' (Scanlon 1998: 150).

Second, Scanlon rejects the idea that justification of norms requires real discourse. Interpersonal interaction 'plays a crucial role in arriving at well-founded moral opinions' (1998: 393 n.5), as understanding generic reasons arising from various standpoints depends in part on the experience of hearing and discussing the claims advanced by people in these standpoints (1998: 395 n.18). But the judgement about what others could or could not reasonably reject 'is a judgement that each of us must make for him- or herself', and thus Scanlon recognises that his account remains 'monological' in Habermas's terms. What Scanlon claims is central to moral motivation is

not the activity of actual justification to others (which does make sense only in relation to individuals with whom we are in contact and communication) but rather the ideal of acting in a way that is justifiable to them, on grounds they could not reasonably reject (1998: 168-69).

So far I have introduced the idea of the democracy/contractualism analogy and argued against the view that Scanlon's contractualism suggests such an analogy. It remains to consider J. Cohen's development of such an analogy, to which I now turn. In this context, the Rousseauian ideas discussed above in relation to Scanlon and Rawls will provide a framework for interpreting J. Cohen's position, as well as for seeing both why it represents an untenable middle ground between Rawls and Habermas and how Habermas develops an alternative procedural approach that can more effectively respond to Estlund's criticisms. I proceed in two stages, first outlining J. Cohen's conception of a democracy/contractualism analogy, and how this relates to Rawls's approach, before arguing that J. Cohen's appeal to Habermasian ideas does little to strengthen Rawls's approach while at the same time giving up one of its most important aspirations.

8.3 Mirroring a Deliberative Initial Situation

If Scanlon does not provide a model of idealized agreement in accordance with an agent-neutral consequentialist grounding of morality, nor a Kantian grounding of morality in universalizability and autonomy, then there is no ideal model for real processes of democratic deliberation to seek to mirror. But this leaves J. Cohen's approach, appealing as is does to Habermas's ideal speech situation, not Scanlon's reasonable rejectability test, untouched, and this could provide an alternative basis for the democracy/contractualism analogy.

The ‘Rawlbermasian’ position that J. Cohen represents retains Rawls’s dualistic architectonic of an idealized initial situation for the derivation of substantive principles and real institutional contexts in which these principles are realized. However, in lifting Rawls’s veil of ignorance it sees itself as rejecting, with Habermas, the ‘Rawlsian attempt to insulate deliberations about principles of justice from the determinate preferences, motives and beliefs of citizens in an original position’. This is seen as making it possible to ‘solve a problem of motivation that emerges when theorists encounter a “rift” or “gap” between what a candidate norm requires and what citizens are motivationally capable of doing’ (Munro 2007: 447). In seeking to close this gap, J. Cohen’s institutional account of the context of realization draws upon Habermas in providing an account of the institutions of deliberative democracy.

J. Cohen derives from Rawls an ideal of democratic politics that involves ‘public deliberation focused on the common good, requires some form of manifest equality among citizens, and shapes the identity and interests of citizens in ways that contribute to the formation of a public conception of common good’ (J. Cohen 1997b: 69). Rawls is seen as offering a formal argument for this model in his argument that the parties in the original position would choose a principle of participation, backed up by a guarantee of the fair value of political liberties. But J. Cohen also sees Rawls as suggesting an informal argument for this conception when he says that:

The guarantee of fair value of the political liberties is included in the first principle of justice because it is essential in order to establish just legislation and also to make sure that the fair political process specified by the constitution is open to everyone on a basis of rough equality. The idea is to incorporate into the basic structure of society an effective political procedure that *mirrors in that structure the fair representation of persons achieved by the original position* (PL: 330, emphasis added).

J. Cohen seeks to further develop this informal argument. Whereas he sees Rawls as deriving principles of justice that constrain democratic procedures, J. Cohen wants to show instead how principles of justice are constitutive of a properly understood democratic ideal: ‘we ought not to proceed by seeking to “mirror” ideal fairness in the fairness of political arrangements, but instead to proceed by seeking to mirror a system of ideal deliberation in social and political institutions’ (1997b: 71).

In developing this model, J. Cohen begins from a formal conception of a deliberative democracy, understood as an ongoing, independent and pluralistic association. Members share a view that deliberation among equals is the basis of legitimacy, recognise one another as having deliberative capacities and see it as important that ‘the terms of their association not merely *be* the result of their

deliberations, but also be *manifest* to them as such' (1997b: 72-73). He then sets out an ideal deliberative procedure to 'give *substance*' to this formal account. This procedure provides a model for institutions to mirror and characterises the 'conditions that should obtain if the social order is to be manifestly regulated by deliberative forums of collective choice' (1997b: 73). Ideal deliberation is free, unconstrained by prior norms, and agreement achieved by deliberation provides in and of itself sufficient reason for complying with that agreement. Ideal deliberation is also reasoned, recognising no force but that of the better argument. Third, the parties are equal, both formally equal in having equal standing and *substantively* equal as a result of the existing distribution of power and resources not affecting their deliberations or their chance to contribute to those deliberations. Finally, ideal deliberation aims to arrive at a rationally motivated consensus.

Estlund argues that Rawls's veil of ignorance makes his original position completely unpromising as a basis for a democracy/contractualism analogy (Estlund 2008: 240). J. Cohen, by contrast, argues that Rawls does suggest such an analogy but does not develop it in detail. For Estlund, the idea of deliberation in the original position is 'removable from the account without essential change' (Estlund 1993b: 1461), while in actual political procedures Estlund thinks Rawls 'argues that voters ought to address justice itself, the primary question the contractual situation seeks to explicate' (Estlund 2008: 240). J. Cohen, on the other hand, argues that Rawls has an informal argument for mirroring a deliberative standpoint in institutions. However, in the light of the argument of the previous two parts there are reasons to believe that Rawls has a more well-developed democracy/contractualism argument than both Estlund and J. Cohen recognise, and one that can provide a basis for assessing their alternative approaches.

Rawls's original position models a procedure of rational deliberation that claims to be congruent with persons' conceptions of the good pursued from the standpoint of deliberative rationality. J. Cohen's post-Rawlsian objections to Rawls's mirroring view relate to the way in which these two standpoints are connected. Rawls's well-ordered society in *TJ*, like Rousseau's account of a well-ordered society, 'has a determinate general will whose content is given by the principles of justice'. While on Rawls's view justice requires political unanimity, J. Cohen argues that 'we cannot expect the most reasonable democratic society to be founded on an agreement about justice' (J. Cohen 2003: 129). Bernard Manin (1987) mounts a similar critique of Rawls's Rousseauian assumptions in a manner that illuminates the strengths and weaknesses of Rawls's model of the relationship between these two deliberative standpoints. In a sense associated with Aristotle, deliberation is the process of the formation of the will in which an agent

assesses different options before deciding on one of them.¹²⁵ Rawls's account of deliberative rationality, and its associated Aristotelian principle, follows this conception. A different sense of deliberation is associated with Rousseau's description of the role of deliberation in the formation of the general will:

From the *deliberations* of a people properly informed, and provided its members *do not have any communication among themselves*, the great number of small differences will always produce a general will and the decision will always be good. But if groups, sectional associations, are formed at the expense of the larger association, the will of each will become general in relation to its own members and private in relation to the state (Rousseau 1997 [1762]: 73, emphases added).

Deliberation is here opposed to communication and equated with decision; Rousseau's position is positively anti-discursive. The reason for this is that 'Rousseau's individuals are already supposed to know what they want when they come to a public assembly to decide in common'. They 'do not deliberate, not even within themselves' (Manin 1987: 346, 347).¹²⁶

All the elements of Rousseau's conception – the requirement for unanimity, the absence of communication and the predetermined will of individuals – are present in Rawls's idea of the original position. The interesting question, though, is how Rawls connects this conception of unanimity with the idea of unanimity in a well-ordered society. Does the appeal to the deliberative standpoint of public reason in *PL* help to resolve the problem that J. Cohen and Manin identify in *TJ*? Rawls takes the idea of public reason to originate with Rousseau's description of deliberative reason as a capacity to adopt the point of view of the general will, and for Rawls the point of view of the general will is represented through the original position (LHPP: 230-31). As Michael Saward has noted, public reason is not public reasoning. This means that rather than Rawls's political turn representing a new departure, a shift from the question of justice to the question of legitimacy, '*the arguments of Political Liberalism can properly be interpreted as the mirror image of those of A Theory of Justice*' (Saward 2002: 120). To be more precise, since the stability argument in *TJ* already contained an account of the deliberative standpoint mirroring the standpoint of the original position, and *PL* retains the two-stage structure of *TJ*, what we see is a shift of emphasis within an existing mirroring architectonic.

¹²⁵ Compare Goodin (2000) on 'democratic deliberation within'.

¹²⁶ While pointing out that Rousseau does not always employ the idea of deliberation in this way, Neuhouser (2008: 202 n.27) broadly accepts Manin's interpretation.

Contrary to what Estlund suggests, Rawlsian citizens do not seek to identify procedure-independent substantive principles of justice. The content of public reason is given by a reasonable political conception of justice, and, when voting, reasonable citizens desire to find a justification for their reasons within the terms of these reasonable principles. Substantive principles are not independent of democratic procedures of citizen deliberation but constitute them. They are not external to citizens' individual conceptions of the good, but rather express the sense of justice with which citizens seek to render their conceptions of the good consistent. Rawls's duty of civility in public reason gives a view of voting that is more than a little 'reminiscent of Rousseau's *Social Contract*' (PL: 219).

There is some irony in the fact that while Rawls is often criticised by self-styled deliberative democrats influenced by a Habermasian conception of discourse for a deficient account of deliberation, Rawls's Rousseauian conception of deliberation is faithful to the republican ideal that Joseph Bessette (1981) coined the term 'deliberative democracy' to describe. If Rawls is not a *discursive* democrat, he is a deliberative democrat *avant la lettre*. 'When the framers rejected direct popular participation in the governing process, they put their faith instead in political institutions. Majority rule properly understood was taken to involve the rule of the deliberative majority' (Bessette 1981). Rawls agrees: 'Constitutional arrangements compel a majority to delay putting its will into effect and force it to make a more considered and *deliberate* decision' (TJ: 229/201, emphasis added). Majority rule is subordinate to the political ends of an 'ideal constitution of public deliberation in matters of justice'.¹²⁷ Thus when Rawls refers to the educational role of 'deliberative political discussion' there is no suggestion of an oxymoron (JF: 146).¹²⁸ Estlund overlooks the republican pedigree of deliberative democracy when he argues that Rawls lacks 'any account of the legitimacy of actual legislative practices in terms of their deliberative nature' (Estlund 1993b: 1462).

8.4 Beyond 'Rawlbermas'

I have sought to clarify the mirroring structure of Rawls's approach, and the Rousseauian ideas it embodies, in relation to J. Cohen's own development of the idea of mirroring a deliberative initial situation in democratic institutions. We are now in a

¹²⁷ For Rawls, while ideally conducted *discussion*, and pooling knowledge in the manner of Condorcet's jury theorem, is more likely to arrive at a correct conclusion than the *deliberations* of any one person, discussion plays only an instrumental role. What guarantees impartiality at the legislative stage of Rawls's four-stage sequence is that legislators are 'constrained by a just constitution' (TJ: 357/314).

¹²⁸ Freeman (2000) also understands deliberative democracy in this sense.

position to assess the extent to which J. Cohen's position represents an advance over Rawls's. Under the influence of a Habermasian idea of discourse, there is tendency to assume that Rawlsian deliberative democracy has an extensive 'external-collective' dimension as well as an 'internal-reflective' one.¹²⁹ This is untrue in the case of Rawls's Rousseauian conception of public reason, and the case is not far different in the case of J. Cohen. In language resonant of Habermas, J. Cohen does seek to 'tie the exercise of power to conditions of public reasoning' in order to establish conditions for discursive will-formation on the part of citizens and generate communicative power (J. Cohen 1998: 186). But for J. Cohen public reasoning is not public discussion (1998: 193). Public reasoning should be directly *institutionally tied* to the exercise of power, as political ideals and values are seen as being acquired in part by participation in common, public institutions, which 'shape the moral-political education of citizens' (1998: 190). Appropriate institutions are needed to '*make deliberation possible*' (J. Cohen 1997b: 79). J. Cohen therefore shares Rawls's intention of embodying the general will in institutions that must 'provide the framework for the formation of the will' (1997b: 79-80). The requirement of shared reasons for the exercise of political power achieves democratic community. It 'expresses the full and equal membership of all in the sovereign body . . . and establishes the common reason and will of that body' (J. Cohen 1998: 222).

An idea of public reason is one element that Rawls takes to be essential to deliberative democracy. Another is that citizens need to have the knowledge and motivation to realize an ideal of public reason in their political conduct. These elements are consistent with J. Cohen's view, as are the main policy implications of Rawls's conception of deliberative democracy: the public financing of elections and civic education. Where J. Cohen departs from Rawls is in a final element. Rawls describes 'a framework of constitutional democratic institutions that specifies the setting for deliberative legislative bodies' (PL: 139). J. Cohen, on the other hand, extends the forums for public deliberation to political parties and secondary associations (J. Cohen 1995). It is in these more extensive institutional provisions that J. Cohen's social theory departs from Rawls's constitutionalism and incorporates a more Habermasian conception of civil society, however limited. There may be more public reasoning, in the sense of interpersonal communication, in this model – the general will may not be immediately apparent as it is in Rawls – but ultimately it is a Rousseauian conception of

¹²⁹ I borrow these terms from Goodin (2003).

the deliberation made possible by institutions to which J. Cohen appeals rather than Habermas's conception of a discursive public sphere.¹³⁰

J. Cohen shares Estlund's definition of, and opposition to, the fair proceduralism of Dahl and Hampshire, but shares Rawls's understanding of the need for, and nature of, substantive principles of justice. For J. Cohen, as for Rawls, 'democracy, on the deliberative conception, is a substantive, not simply procedural, ideal and . . . the substance comprises egalitarian and liberal political values' (J. Cohen 1998: 187).¹³¹ Specifically, this substance, involving principles of deliberative inclusion, the common good, and participation, includes values of equality and liberty (J. Cohen 1996: 113).¹³² A deliberative conception of democracy, unlike an aggregative conception, 'does not face the same troubles about reconciling democracy with non-political liberties and other substantive, nonprocedural requirements' (1996: 99). J. Cohen follows Scanlon's substantive account of moral motivation, but does not seek to mirror a 'Scanlonian original position', focusing like Rawls on a background of cooperative institutions that is only implicit in Scanlon and 'constrains what can count as an acceptable reason within a process of deliberation' (1996: 101). Herein lies J. Cohen's basic commitment to liberal constitutionalism in contrast to Habermas's discursive Kantian Republicanism.

While the limitations of Rawls's conception of democracy have frequently been remarked upon, their strengths have not. Rawls's democratic ideal of a well-ordered society may be overly substantive – too Rousseauian and Hegelian and insufficiently Kantian – but its republican ideal of freedom, according to which citizens regard their political arrangements as products of their will rather than constraints upon it, is attractive. Any plausible post-Rawlsian approach should seek to retain this. J. Cohen's 'Rawlbermasian' position integrates a conception of deliberative democracy inspired by Habermas into what remains an essentially Rawlsian liberal constitutionalist framework.¹³³ However, J. Cohen's account lacks the resources of Rawls's conception of democracy as a relation of citizens within the basic structure. While the weakness of this conception is its substantive assumption of a shared social world, its strength is its conception of how in mutually acknowledging the great value of the social world they share, citizens experience popular sovereignty as an expression of their collective will, the will of the people. J. Cohen's position marks little advance over Rawls's. Crucially lacking is an equivalent to Rawls's conception of reflective equilibrium, leaving J.

¹³⁰ This is not altered by the conception of directly-deliberative polyarchy in J. Cohen (1997).

¹³¹ See J. Cohen (1991) for a critique of Dahl, and J. Cohen (1994a) for a critique of Hampshire.

¹³² J. Cohen (1998: 207) presents this as a shift from J. Cohen (1997b), but it is not inconsistent with it.

¹³³ Saward (2000: 75) describes J. Cohen's approach as 'non-deliberative – possibly anti-deliberative'.

Cohen unable to theorize how ideal principles are mirrored, in the sense of being implicit within, real social practices. J. Cohen neither constructs an ideal that seeks to find its reflection in a real deliberative standpoint, like Rawls, nor reconstructs the ideal presuppositions of real discourse, like Habermas. In extending deliberative democracy beyond Rawls's public political forum, J. Cohen does not reflect the concern that drives Rawls's division of justificatory labour to incorporate a standpoint from which citizens can reflectively understand democratic arrangements as authorised by their will. There is a 'hasty dive from the great heights of the theory to the advocacy of deliberation in political parties and secondary associations' (Saward 2000: 75), with little conception of how these are related.

Ultimately the differences between J. Cohen and Rawls are more semantic than substantive. J. Cohen has insightfully elaborated what Rawls means by saying that justice as fairness is 'for a democratic society' (J. Cohen 2003, 1994a). But in the idea of public reason Rawls develops his own vocabulary for showing how the terms of the original position can be regarded as the regulative framework required to realize the values of democratic community. Rawls moves from substantive principles of justice to an increasing concern with how they embody democratic values, while J. Cohen converges on a similar framework by beginning from democratic values and then arguing for substantive principles as the precondition for realizing them. While J. Cohen may be right that contemporary societies lack the general will necessary for a consensus on substantive principles of justice, his alternative of a consensus on a substantive conception of democracy is an attempt to regenerate this general will without accounting for the motivational basis upon which citizens could endorse such a conception as expressive of their will.

8.5 Rescuing Proceduralism

Estlund is therefore right to suggest that J. Cohen lacks an account of the normative status of rights (see also Cooke 2002: 84). It is not so much that the idea of democracy 'is being stretched too thin to be recognizable', but that substantive rights are definitionally made part of the idea of democracy rather than being seen as a constitutive element. But I do not follow Estlund in believing that '[a] principle of democracy must surely assert something like inalienable popular sovereignty', in the sense of making democracy equivalent to a majoritarian process (Estlund 2008: 92).

If J. Cohen's substantive 'Rawlbermasian' position represents little advance over Rawls, then two options are left open, either following Estlund's epistemic

proceduralism or Habermas's discursive proceduralism. I will argue for the latter course in the remainder of this thesis, and this argument begins in this section through a critique of the way Estlund seeks to justify his epistemic proceduralism in relation to the positions of Rawls and Habermas.

Estlund's model of proceduralism shares the methodological individualism of social choice theory. For Estlund, J. Cohen's account is procedural not substantive as 'once all the democratically motivated *constraints* [of democratic procedures] are fully respected, the value of democratic procedures is held to be intrinsic in some way and not based on any tendency to *promote* other values such as justice or the common good' (2008: 91, emphases added). As with G.A. Cohen's critique of Rawls, institutions are seen as constraints on the promotion of the ends of justice. As we have seen, Estlund wants to steer a middle course between proceduralism and a Platonic correctness view. Alongside Plato, Estlund also sees Rousseau as committed to a correctness view, but this is because he interprets Rousseau from the epistemic perspective of Condorcet's jury theorem. Although he rejects the specifics of Condorcet's jury theorem, it provides the inspiration for Estlund's own epistemic defence of democracy.¹³⁴

I earlier presented Estlund's epistemic proceduralism as a model of constrained optimisation. We can analyse Estlund's position by looking at how it develops a conception of the constraints of democratic legitimacy in relation to Rawls and a conception of epistemic optimisation of justice in relation to Habermas.

First, then, the conception of legitimacy. In forestalling the move towards a correctness view, Estlund's procedural constraint of legitimacy adapts Rawls's idea of reasonableness to give a qualified acceptability requirement. In doing so, Estlund (1998a) rejects Rawls's reticence to appeal to the truth, arguing that a qualified acceptability requirement must itself be accepted as true. Qualified acceptability is reflexive: it applies to itself. 'It says that political justifications cannot appeal to doctrines that are not acceptable to all qualified points of view, and it is itself a doctrine appealed to in political justification' (Estlund 2008: 53). Estlund's acceptability requirement is the following:

AN (for 'acceptance necessary'): No doctrine is admissible as a premise in any stage of political justification unless it is acceptable to a certain range of (real or hypothetical) citizens, C, and no one else's acceptance is required (2008: 53).

¹³⁴ Estlund (1989) acknowledges that it is likely Condorcet was influenced by Rousseau rather than influencing him. Schwartzberg (2008) emphasises Rousseau's concern to elicit 'unanimity for the right reasons'.

Estlund argues that to avoid being self-excluding, AN must specify C so that its members accept that specification. This is required to penetrate the plurality of insular groups, groups which recognise the rejection rights of all and only their own members. According to Estlund, there needs to be an authoritative group. We cannot just get by with ‘we the reasonable’: the authoritative group will have to say that other groups are *mistaken*.

However, for Rawls the idea of reasonable citizens does not demarcate a class of citizens who hold reasonable views, but rather characterises a form of motivation possessed by citizens within the social world formed by basic institutions. The extent of this class of citizens is shown by the preparedness of citizens to endorse a reasonable political conception of justice for the basic structure in an overlapping consensus. Reflexivity enters into Rawls’s model through the fact that reasonable persons come to recognise the reasonable conception of social unity that they share with other reasonable citizens by reflecting on the congruence of a political conception of justice with their comprehensive doctrines, and then embedding this conception in their comprehensive doctrine and recognising that others have done the same in wide and general reflective equilibrium. For Estlund, qualified persons must accept certain substantive principles – fundamentally they must accept a certain view of who counts as qualified. However, while this assumption is necessary for Estlund’s theory, he does not lay out a substantive criterion for which points of view count as qualified (2008: 63).

One gets a sense of how different Estlund’s concerns are to Rawls’s when Estlund says that he avoids the term ‘reasonable’ as in ordinary language reasonable people would not be guaranteed to accept his own arcane view about qualified acceptability and legitimacy. The term qualified, by contrast, is ‘a term of art for a certain theoretical role’ (2008: 61). But it is essential to Rawls’s view of reasonable moral motivation that it draws on the connotations of ordinary moral language.¹³⁵ Estlund stretches the meaning of morality very thin when he presents epistemic proceduralism as giving citizens ‘moral reasons to comply, not epistemic reasons to believe’ (2008: 106). This is because the moral disposition of citizens so defined is given by a disposition towards truth, including the truth about qualification required to get the theory off the ground. Estlund’s positing of an external relation between Rawls’s reasonable principles of justice and the reasonable moral motivation that on Rawls’s view leads citizens to endorse such principles leads Estlund to interpret both the original position and an overlapping consensus as objectivist procedures that are ‘epistemic of truth’ rather than constitutive of normative rightness (Estlund 1993a: 97 n.20). Rawls’s theory of justice becomes ‘an

¹³⁵ As it is in Sibley (1953), the source of Rawls’s distinction (JF: 7 n.6).

example of how there might be independent standards of the correctness of democratic political decisions' (Estlund 1993b: 1463). But for Rawls, reasonable principles of justice are those that can be justified according to a procedure that incorporates a *moral* conception of citizens as free and equal rather than one that citizens can accept for its *epistemic* benefits in tracking substantive truths about justice. Constructivism, in both its Rawlsian contractarian and Scanlonian contractualist variants, is concerned with unanimity for the right reasons (Freeman 2006).

We have seen how Estlund formulates a legitimacy constraint in dialogue with Rawls. It is now time to look how, in relation to Habermas, Estlund develops the idea that democratic procedures can come to 'know the truth'. Estlund's rejection of the democracy/contractualism analogy is the springboard from which he develops his own 'epistemic departure view', although before he gets there he considers, only to finally reject, a more promising version of the analogy, an epistemic 'mirroring view' (Figure 5).

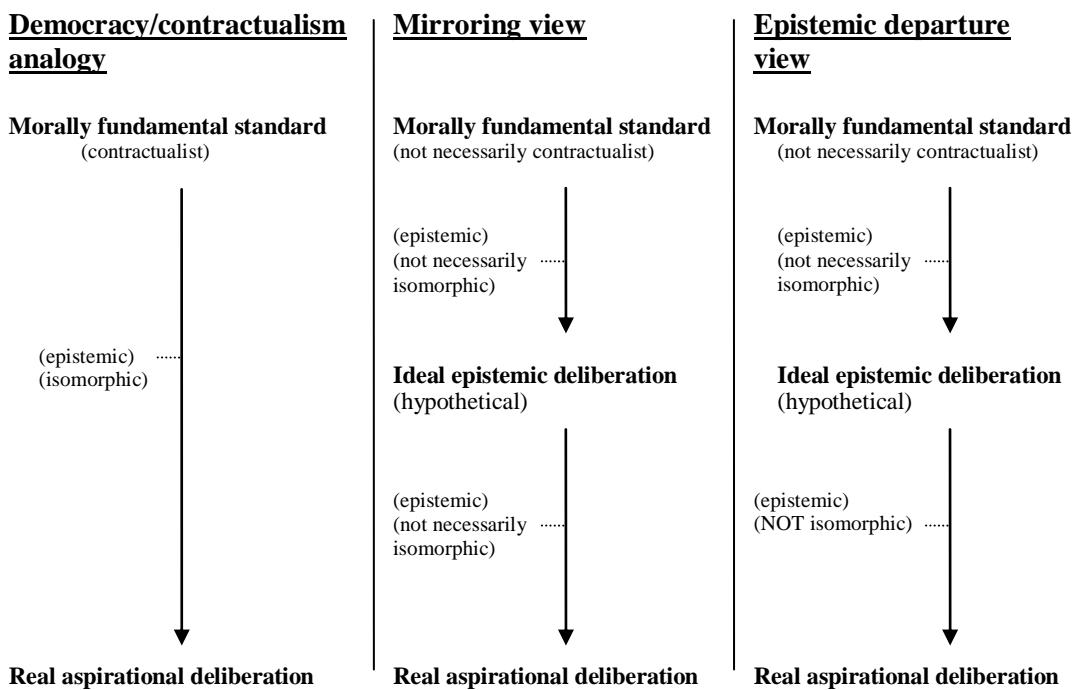


Figure 5. Models of the Relationship between Democratic Deliberation and Justice

Source: Adapted from Estlund (2008: 174).

Estlund's objection to the democracy/contractualism analogy is that participants in deliberation, like contracting parties, cannot be thought of as addressing the primary question 'What is justice?' Instead, they must address less impartial reasons for their

choices, which entitles them to employ a personal veto right that is inappropriate as something to replicate in democratic procedures. The mirroring view, by contrast, seeks to replicate an ideal epistemic deliberation in which deliberators address and debate the primary question of justice (2008: 172). The tracking claim that democratic decisions tend to be just or right by independent standards is supported by a stronger isomorphism as the primary question is addressed in both ideal and real deliberations; neither requires unanimity and in neither do deliberators see themselves as free of all prior normative constraints. Estlund's favoured epistemic departure view is similar to the mirroring view, but it rejects the aspiration for isomorphism between ideal epistemic deliberation and real aspirational deliberation. '[A]ctual practice that is possible and hoped for, is not meant to approximate the ideal epistemic level, much less the morally fundamental standards' (2008: 173).

In setting out this view, Estlund defends a 'real speech situation' that he contrasts with Habermas's ideal speech situation. Estlund recognises that it is a misreading of Habermas to see him as advocating institutions of public discussion that 'approximate' an ideal speech situation. (2008: 199-200). But while politics 'is not meant to resemble such an imaginary situation . . . it ought to have a tendency to produce the *same decisions*' (2008: 89, emphasis added). Thus, as in his critique of J. Cohen, Estlund takes Habermas to tacitly rely upon substantive procedure-independent standards. But Estlund mischaracterises the ideal speech situation in the same manner as Rawls's original position and overlapping consensus by presenting it as epistemic rather than constitutive of moral rightness. While in interpreting Rawls, Estlund separates the idea of justice from the procedure through which principles of justice are constructed, in interpreting Habermas, Estlund rejects the idea that ideal standards are internal to real practices of discourse.

Of course, with respect to Habermas, Estlund's aim 'is not at all exegetical' (Estlund 2008: 185). He recognises that it is ideal practical deliberation that should play a role in politics, not the ideal speech situation, but says that 'the "ideal speech situation" is an evocative phrase that has caught on', and we can 'safely treat it as the overarching idea that unifies Habermas's approaches to descriptive and normative validity' (2008: 184). Where this is problematic, however, is in obscuring Habermas's distinction between descriptive and normative validity. Estlund recognises that while Habermas holds a cognitivist view of morality, he does not hold that moral judgements can be true or false in the same way as descriptive statements, but Estlund argues that Habermas lacks a basis for making this distinction.

Habermas distinguishes between truth and normative rightness in the following terms. The second-order proposition of the normative claim 'one ought to lie' is 'it is

right that [one lies]'. But one cannot make a similar move from 'the table is yellow' to a second-order proposition 'it is yellow that . . .', only to 'it is true that [the table is yellow]'. However, in Moorean intuitionist fashion, Estlund argues Habermas has done nothing to block the move to a third-order predicate, such as 'it is true that, it is right that [one lies]' (1993a: 77).

The problem is that this ignores the constructivist insight that for a proposition to be right it must be understood as right for the right reasons. Now, as we saw in relation to the Rousseauian lawgiver problem, Rawls may not have successfully moved beyond a dualistic model that presupposes a perspective beyond the existing rules of social practices. But moving to the Platonic paradigm is not the only alternative. In the next chapter I explore how Habermas's reconstructive procedural account of moral rightness blocks this move to a third-order predicate. For now, though, we may note that what separates Estlund's model of deliberative democracy from the positions of Rawls and Habermas is Estlund's rejection of a self-legislating conception of practical reason. Estlund equates practical reason with reasons for choice in such a way that Gauthier's model of instrumental reason qualifies as a conception of practical reason. Contrasting with the reasons for choice of practical reason are not theoretical reasons associated with given objects, but rather cognitive reasons to judge or believe that something is so (Estlund 1993b: 1450). Cognitive judgements are in turn distinguished from performative judgements, which seek to fix or determine a situation. Performative judgements may be based on cognitive judgements that can be true or false, but the act of performative judging is not itself true or false. According to Estlund, 'deliberative politics must involve the use of cognitive reason – judgements that can be true or false' (1993b: 1451).

Estlund recognises no intermediate position between instrumental self-interested contractarianism and appeal to external standards of truth. As a result, Estlund's characterisation of proceduralist deliberative democracy is like Ridge's characterisation of contractualism. In order to respond to Plato's Euthyphro question, 'contractualist theories of right and justice cannot, on pain of circularity have participants armed with views about justice or rightness' (Estlund 2008: 172). But for both Rawls and Scanlon, in the procedure of reflecting on our views about justice on the basis of views about rightness that we already hold, we find ourselves in an interpretative and not a vicious circle. While Rawls is unable finally to justify the authority of the standpoint from which substantive principles of justice for the basic structure are derived, and the critical potential of Scanlon's contractualism is blunted by its lack of context-transcending standpoint, Habermas offers the possibility of overcoming this dualism through a

reconstruction of the context-transcending ideal standards immanent in real procedures of discourse.

Conclusion

This chapter has explored the contrasting methodologies of G.A. Cohen, Rawls and Habermas in the context of the development of alternative approaches towards deliberative democracy. I have argued that Estlund's reason for rejecting the practical goal of mirroring a model of ideal deliberation in democratic practice rests on an understanding he shares with G.A. Cohen; it is an appeal to 'aspirational theory' seen as required to combat the 'utopophobia' of complacent realism (Estlund 2008: 259). For Estlund, a standard that demands that citizens address the truth about justice 'should not be removed from normative political theory just because it will not be met' (2008: 275).

But while Estlund's dichotomy of utopian standards of truth and a real speech situation of democratic deliberation is as unacceptable as G.A. Cohen's position, J. Cohen's 'Rawlbermasian' position lacks the resources for showing why this is so. J. Cohen's position is not made untenable simply by the fact that he does not want to subscribe to Habermas's wider programme (J. Cohen 1999) and conception of the standards internal to discourse (J. Cohen 1998: 198). The fundamental problem is that in rejecting G.A. Cohen's appeal to fact-independent principles by emphasising the Rousseauianism and Quinean holism of Rawls's position (J. Cohen 2002), he makes no advance over Rawls's attempt at reconciling substantive principles and the autonomy of citizens in an approach to political philosophy that aspires to provide the common ground for citizens in a democratic society. The option of exploring the implications of taking onboard Habermas's wider programme therefore grows in attractiveness.

There are two leading explanations of how Rousseau sought to overcome the problem of utopianism: the figure of the Lawgiver and a Condorcetian epistemic conception of voting (Honig 2007). I have rejected Estlund's epistemic democracy, which, while drawing on Rousseau, reads him within G.A. Cohen's Platonic paradigm, and J. Cohen's approach, which does not resolve the problem of the Rousseauian political circle that bedevils Rawls's approach. In the next chapter I turn to Habermas's reconstructive utopian realism in order to assess whether in its alternative elaboration of the Rousseauian paradigm it can avoid the problem of utopianism.

PART V. HABERMAS'S UTOPIAN REALISM

9. Reconstructivism and Transcendence from Within

In seeking a perspective from which to view Habermas's project as a whole, *BFN* provides an appropriate vantage point. Its methodological attempt to mediate the tension between facticity and validity expresses a concern for the intrinsic relation of history to reason that goes back to Habermas's dissertation on Schelling (Dews 2001: 205-209). Habermas does not oppose the ideal to the real, but rather attempts to reconstruct, independently of a philosophy of history, 'particles and fragments of an "existing reason" already incorporated in political practices, however distorted these may be' (BFN: 287). In its political focus, *BFN* returns to themes Habermas first addressed in *The Structural Transformation of the Public Sphere*, articulating a theory of law that integrates the discourse theory developed in discourse ethics into Habermas's two-level social theory of system and lifeworld set out in *The Theory of Communicative Action*.

Started before and published three years after the events of 1989, the achievements of the Federal Republic of Germany and the challenges of the transition process of reunification lend an important impetus, even if Anglo-American literature often provides Habermas's frame of reference (Specter 2009). Habermas asks how the socialist project of radical democracy can be reconceived as an ongoing project – as opposed to the design and implementation of a concrete form of life (Habermas 1996b: 478; BFN: xli) – in a manner that takes account of the circumstances of a complex modern society. As always, 'modernity, now aware of its contingencies, depends all the more on a procedural reason, that is, on a reason that puts itself on trial . . . there is neither a higher nor a deeper reality to which we could appeal' (BFN: xli). In placing the problem of self-grounding centre stage, I side with the minority view that neither Habermas's engagement with the problem of social complexity nor his response to the collapse of actually existing socialism represents an abandonment of the project of radical democracy. We should take seriously Habermas's claim that:

A world-historical event like the collapse of the Soviet empire certainly requires us to rethink our political positions, but for decades I have continued to defend a radical reformist line. Despite all the changes in my theoretical position, I also understand the discourse theory of law in a radical democratic sense (RSP: 442; see also Power 1998; Grodnick 2005; McCarthy 2003).

The book's German title – *Faktizität und Geltung: Beiträge zur Diskurstheorie des Rechts und des demokratischen Rechtsstaats* – brings Habermas's concerns into

focus. Habermas employs discourse theory to conceptualise the relationship between law, and legal rights, and the democratic law-governed constitutional state, seeking to work up the ‘hunch’ that postmetaphysical law depends upon radical democracy into insights into a conceptual or internal relation between the rule of law and democracy (BFN: xlii, 449). Like Rawls, but with a focus on positive law, Habermas seeks to mediate between Platonism and positivism, Habermas arguing against both the idea that positive law can draw its legitimacy from a higher law and the legal positivist view which denies that there is anything to the legitimacy of law beyond the contingency of legislative decisions (BFN: 450-53). Methodologically, this requires integrating the perspectives of law as social fact and as legitimate norm through focusing on the tension between facticity and validity within the medium of law. In a reciprocal relation, ‘[t]he theory of communicative action already absorbs the tension between facticity and validity into its fundamental concepts’ (BFN: 8), while *BFN* theorizes how the tension between facticity and validity in communicative action can become politically effective through being translated into the medium of law.

Habermas’s study proceeds through a social-theoretical analysis of the emergence of modern law and its role in contemporary societies, a legal theory that reconstructs the relationship between the rule of law and democracy, a jurisprudential discussion that applies discourse theory to the rationality of legal adjudication, and a sociological conception of deliberative politics that integrates normative and empirical elements. In this chapter I focus on these initial two stages, while I briefly touch upon the latter two in the next. Since my concern is with how Habermas can be viewed as reconstructing Rawls’s approach, this will of necessity be a summary interpretation. I begin by presenting how Habermas seeks to mediate between Rawls’s philosophy of justice and Luhmann’s sociological systems theory of law. I then examine the roots of the tension between facticity and validity in Habermas’s theory of communicative reason, presenting Habermas’s analysis of linguistically mediated social practices as the basis on which he defends Kantian conclusions by Hegelian means. An important excursus follows in which I justify my focus on Habermas’s political theory by clarifying its relationship to his discourse theories of truth, morality and ethics. Finally, I assess Habermas’s reconstruction of the relationship between law and democracy.

9.1 Beyond Rawls and Luhmann

Habermas begins with the internal tension between facticity and validity in communicative reason and the law, but the rationale for this starting point becomes

clearer when one first considers the external tension between facticity and validity with respect to law. Here social facticity consists in the facts of social complexity that constrain the operation of the legal system (BFN: 34, 288). Habermas seeks to mediate between the disciplinary perspectives of the sociology of law, represented by Luhmann, and the philosophy of justice, represented by Rawls, a contemporary reincarnation of a dispute between a Hobbesian approach that focuses on positive law to the exclusion of morality and a Kantian approach that subordinates law to morality (BFN: 270). The concepts of a purely normative philosophy of justice will be empty without an understanding of law as an empirical action system, while a sociological perspective that lacks hermeneutical access to internal meanings is blind (BFN: 66). It follows that a reconstruction of the *internal* tension between facticity and validity within law will have to make sense of the twin perspectives from which addressees of law can view a legal norm: the performative attitude in which a legal norm claims normative validity and is followed out of a sense of duty and the objectivating attitude in which a legal norm is regarded as a coercive social fact (BFN: 30, 114).

From Habermas's perspective, the shortcomings of Rawls's attempt to 'bridge the chasm between ideal theoretical demands and social facts' is that the resistant reality with which critical reason wants to keep in touch is not primarily a culture of reasonable pluralism but 'the harder material of institutions and action systems' (BFN: 64). This shows the extent to which Habermas has both learned from Luhmann's systems theory as well as criticising it.¹³⁶ Luhmann is part of a tradition of realist objections to the rational natural law view of society as a free association of autonomous and equal citizens. Adopting an observer's perspective, Marx sought to expose the ideological nature of this claim, but in eliminating the Marxian conception of teleological development, Luhmann radicalises a realist perspective, conceiving of society as an autopoietic (self-creating) self-referential system. Modelled on cybernetic systems, autopoiesis means that the legal system describes itself in legal terms, thereby constituting and reproducing legal acts by its own means. The sociological observer is self-referentially subsumed within the system he observes, viewing himself and his science as just one subsystem among others (BFN: 47-49). In contrast to Rawls's view of an external philosophical standpoint becoming integrated into a spiral of social learning, here we have a self-perpetuating spiral through which law insulates itself against assessment from an external normative perspective.

Luhmann's social theory provides a telling corrective to Rawls's social theory, and Habermas's criticisms here are more perspicuous than in the Rawls-Habermas

¹³⁶ See Habermas (1979b) for his early engagement. Luhmann (2002) is a polemical summary of his objections to Habermas.

debate where they tend to be obscured by objections from the perspective of discourse ethics. However, it is still the case that Habermas fails adequately to represent Rawls's own attempt to reconcile morality and politics. Putting Habermas's criticism of Rawls's 'bare normativity' together with his criticism of the way in which Rawls's concern with stability compromises the idea of normative validity lends Habermas's overall judgement of Rawls something of a 'damned if he does and damned if he doesn't' character. Habermas takes Rawls's treatment of the problem of stability in *PL* to be ambiguous between pursuing at a deeper level the problem of social self-stabilisation already treated in *TJ* and seeking to answer the 'question of how one can, in the present circumstances of a democratic society, see to it that the theory finds the measure of well-considered acceptance necessary for a reformist improvement of existing institutions in light of the theory' (BFN: 60). Habermas follows the second reading, with the result that he criticises Rawls both for compromising the claim to universality of moral validity and for a one-sided normative detachment from social facts. The rationale of Rawls's project fails to emerge because Habermas does not consider a third reading according to which *PL* develops the concern for stability in *TJ* in the direction of a realistic utopianism.

As in the Rawls-Habermas debate, Habermas comes close to recognising the true nature of Rawls's position when he observes that Rawls's political conception of justice responds to a Hegelian problem. In Rawls's well-ordered society, 'the individual's morality (*Moralität*) would find its ethical (*sittliche*) context in the institutions of a just society' (BFN: 58). Habermas's rationale for relying on a Kantian theory of law is that Hegel's model 'sets unattainable standards for us' (BFN: xxxix). All attempts to approach social practices from a purely normative perspective represent an outmoded natural law tradition that 'up to Hegel wanted to single out normatively the only reasonable social and political order' (BFN: 3). But for reasons that are understandable given his programmatic intent, Habermas does not consider whether Rawls aspires to albeit more modest but still recognisably Hegelian standards. Habermas also downplays the extent of his own allegiance to such standards. When Habermas claims 'What could once be coherently embraced in the concepts of Hegelian philosophy now demands a pluralistic approach that combines the perspectives of moral theory, social theory, legal theory, and the sociology and history of law' (BFN: xxxix), it is evident that the interdisciplinary approach he proposes retains the ambition to coherently embrace the traditional aspirations of *Rechtsphilosophie* on a new basis.¹³⁷

¹³⁷ 'Like Hegel, Habermas seeks to fashion a philosophy of right, or law, which surmounts the oppositions of empirical and normative considerations, of reason and reality, philosophical right and positive law, and facticity and validity' (Buchwalter 2002: 129). See also Howard (1996).

Without considering the change in perspective of the paradigm shift to a philosophy of language, Habermas's equation of Rawls's notion of reflective equilibrium with his own understanding of reconstruction at the two stages of first justification and then stability is misleading. Habermas complains that Rawls blurs the distinction between the context of justification before philosophical experts and the context of advocacy before fellow citizens (BFN: 59-60). But Rawls, a self-styled student of philosophy writing as a citizen among other citizens, would refuse this distinction.

The apparent contradiction in Habermas's criticisms are dissolved when one recognises that Habermas is arguing that philosophy, which has become a specialized activity unable to provide guidance for a form of life as a whole, must become more abstract in order to remain in touch with social reality. What Habermas means by maintaining contact with reality is maintaining contact with reality as studied by the social sciences, the reality of a society in which philosophy can no longer seek an unmediated contact with society. Rawls's social theory views social institutions purely from the functional perspective of how they both make possible and are amenable to regulation by a reasonable political conception of justice. Rawls regards social institutions solely from the perspective of stabilising a tension between facticity and validity and not from the perspective of a resistant reality programmed by its own logic and immunized against normative inputs. The tension between facticity and validity in individuals' attempts to achieve reflective equilibrium is stabilised within the formation of an overlapping consensus of reasonable comprehensive doctrines within the basic structure of society in a hoped for reconciliation to our social world.

Luhmann's systems theory can claim to have resolved the relationship between theory and practice left unresolved by Hegel and Rawls as 'by insisting on the reflexivity of all structures, Luhmann incorporates his own activity into his functionalist model and thereby establishes a unity of theory and practice' (Holub 1991: 110). But this resolution of the problem of self-grounding sees social complexity as rendering Enlightenment ideals obsolete. Habermas's strategy is to focus on the tenability of Luhmann's claim to be observing the system of law from a neutral, value-free perspective without reference to the performative perspective of the addressees of law. To describe Habermas's strategy as working *between* Rawls and Luhmann (Rehg 1996: xx) gives the same misleading impression as the idea of occupying a middle ground *between* facts and norms. In reconstructing how facticity and validity interpenetrate *within* the medium of law, Habermas is pursuing an approach which can claim to overcome the opposing positions of Rawls and Luhmann. With a sociological analysis of law that combines external access with an internal reconstruction it ceases 'to be

necessary for normative theory to seek contact with social reality in an unmediated way, through the political consciousness of a public of citizens' (BFN: 65).

While the following remark of Habermas's gives no credence to Rawls's post-Kantian Hegelianism, it pinpoints a real weakness in Rawls's position that Habermas can claim to be able to remedy:

In his reflections on the gap between the is and the ought, John Rawls distinguished between 'ideal' and 'real' theory [sic.]. This methodological distinction did not go far enough in detranscendentalizing the Kantian distinction between the world of noumena and the world of phenomena. Ideas find their way into social reality through the unavoidable *idealizing presuppositions of everyday practices* and in this way inconspicuously acquire the status of stubborn facts (BNR: 334).

Detranscendentalization must move beyond not only Kant's dualism of noumenal and phenomenal worlds, but Rawls's dualism of constitutive rules for the basic structure that are regulative of our sense of justice, moving from the level of ideas that form the basis of the social practice of the basic structure to the more fundamental level of the social practice of communication. In reconceiving of practical reason in terms of communicative reason, practical reason 'no longer provides a direct blueprint for a normative theory of law and morality' (BFN: 5). Instead:

A set of unavoidable idealizations forms the counterfactual basis of an actual practice of reaching understanding, a practice that can critically turn against its own results and thus *transcend* itself. Thus the tension between idea and reality breaks into the very facticity of linguistically structured forms of life. Everyday communicative practice overtaxes itself with its idealizing presuppositions, but only in the light of this innerwordly transcendence can learning processes take place at all (BFN: 4-5).

Evidently Habermas's political theory puts to work a network of conceptual resources developed in earlier work; the task of the next two sections will be to examine the most salient.

9.2 Communicative Reason and Reflexive Social Practices

A reconstruction of Habermas's project may begin with the concept of facticity. Literally facticity refers to the quality or condition of being a fact. Habermas describes the 'facticity of facts' in his critique of positivism:

Facts are given evidently in sense experience. Simultaneously they have the immovability and indisputability of something given intersubjectively. The facticity of the facts is evidence both of the certainty of subjective perception and of the external existence of a matter of fact, mandatory for all subjects (KHI: 81-82).

In conceiving of knowledge in scientific terms as a copy of reality, positivism seeks to eliminate the metaphysical idea of the ‘thing-in-itself’ lying beyond appearances, but succeeds only in blocking off reflection on its own metaphysical presuppositions. Positivism has no notion of the way in which the knowing subject provides the subjective conditions of the objectivity of possible knowledge, and, consequently, ‘every conceivable form of transcendence is retracted’. ‘In its unconcealedness, facticity knows no opposition of essence and appearance, of being and illusion, because the facts themselves have been elevated to the status of essence’ (KHI: 85). Habermas no longer makes such points in the language of Hegelian-Marxism, but the lesson remains the same. In drawing attention to the resistance or facticity through which facts exist as facts, Habermas highlights how facts are not copies of external objects but exist through active engagement with objects (TaJ: 27).

If opposition to positivism remains a constant, Habermas’s methodological approach undergoes an important shift in the move from the philosophy of the subject to the philosophy of language. Self-reflection in the framework of ideology critique is abandoned in favour of reconstructive sciences that ‘*systematically reconstruct the intuitive knowledge of competent subjects*’ (OPC: 29). Reconstructivism makes explicit implicit knowledge; in Gilbert Ryle’s terms it transforms know-how into a second-level know-that (OPC: 33; see also TCA 1: 8). As McCarthy explains the contrast:

Whereas critique is brought to bear on something particular, reconstruction deals with anonymous rule systems that any subject may follow insofar as he has acquired the requisite competences. Whereas critical self-reflection makes unconscious factors conscious in a way that has practical consequences, reconstruction renders explicit a know-how (that is, the intuitive knowledge acquired with a competence) without involving practical consequences of this sort (McCarthy 1978: 101).

While critique is therefore context bound, directed at rendering transparent totalities (PDM: 300), reconstruction identifies a potential that can transcend its historical context.

Habermas takes on board and develops lessons from the philosophy of language in engaging with the methodology of the social sciences. Following the conclusions of Chapter 7, particularly relevant in the present context is how Habermas relates his position to Wittgenstein’s model of rules and Davidson’s critique of Quine’s scheme-content dualism. Habermas recounts the evolution of linguistic analysis through three

stages of self-reflection: the transcendental stage of the early Wittgenstein, the sociolinguistic stage of the late Wittgenstein and the stage of historical reflection reached by Gadamer (LSS: 171). On a transcendental conception, reflection on ordinary language seeks to step outside ordinary language. Noam Chomsky is presented as representative of this view. It is interesting to note that while Rawls introduces the role of reflective equilibrium in describing our sense of justice by analogy with Chomsky's idea of the sense of grammaticalness we have for the sentences of our native language (TJ: 47/41), he seeks to bring the observer's standpoint within a social context, even if the problem of unifying the language games of different comprehensive doctrines is ultimately only resolved by assuming a perspective beyond them. Rawls therefore reaches something akin to the sociolinguistic stage. But Gadamer's hermeneutics, instead of seeking a meta-theory, seeks to preserve 'the unity of reason in the pluralism of languages' by appealing to 'the tendency to self-transcendence that is inherent in the practice of language' (LSS: 144). What Gadamer sees as essential to language is the ever-present possibility of translation: in learning our native language we also learn how to learn languages in general; in learning a rule we learn how to interpret and not merely apply that rule. Methodologically, the theorist as interpreter is a moment in the same context of tradition as his object, while that object itself is partners in dialogue who are themselves in the position of interpreters.

Gadamer's approach has similarities with Davidson's, but Gadamer's participant's perspective opens up insights into shared normative standards of rationality that are blocked by Davidson's observer's perspective (BNR: 66). In confining himself to the given objective world, Davidson does not understand sociality 'from the perspective of a participant who "finds himself" within a form of life shared with others, and hence is not only objectively equipped with similar behavioural dispositions but also has at least intuitive awareness of this agreement' (BNR: 64; TaJ: 69). It remains to be seen, though, how Habermas can give Gadamer's hermeneutics a critical dimension that comports more favourably with the philosophical discourse of modernity. To be sure, in Gadamer's understanding of *Verstehen*, 'reflection is no longer blinded by the illusion of an absolute, self-grounded autonomy, and it does not detach itself from the ground of the contingent on which it finds itself' (LSS: 168), but he fails to recognise how linguistically-mediated understanding both shakes and reconstructs, rather than simply reaffirming, the certainties of social practices.

While these methodological reflections precede Habermas's linguistic turn, they illustrate the context out of which it will develop and explain why Habermas describes

his formal pragmatics as equivalent to a transcendental hermeneutics (TaJ: 74).¹³⁸ Habermas's formal pragmatics (originally 'universal pragmatics') is sometimes presented as though it were the context-transcending basis of Habermas's critical theory, albeit one that Habermas seeks to justify on non-foundationalist premises (Cooke 2006: 47-59). However, this fails to acknowledge the way formal pragmatics builds upon a Gadamerian conception of linguistic intersubjectivity, and how this conception in turn informs Habermas's approach to moral and political theory.¹³⁹

The task of formal pragmatics is to 'identify and reconstruct universal conditions of possible mutual understanding' (OPC: 21). The key phenomenon to be explained is the 'peculiar reflexivity of natural languages' (OPSI: 73). Language has not just a propositional but also a pragmatic meaning: we do not just *say* things *with* language but also in *using* language we *do* things. As a result, language is self-explicating or self-interpreting, evident in the fact that competent speakers can paraphrase, or translate, any expression of a language in that language itself. In a model of reaching understanding based upon interpretation, the inherent reflexivity of language is shown in the way language self-referentially combines communication of a content with communication about the sense in which the communicated content is used (OPC: 64).¹⁴⁰ '*[I]n performing speech acts we also performatively produce the conditions that make possible the utterance of sentences in the first place*' (OPSI: 73). Habermas appeals to speech act theory and Austin's idea of 'how we do things with words' to explicate the double structure of speech that underlies the reflexive character of language. As Barbara Fultner summarises it: 'Every speech act takes the form *Mp*, where *M* expresses the illocutionary force of the utterance (the communicative dimension) and *p* expresses its propositional content (the cognitive dimension) about which mutual understanding is to be reached' (Fultner 2001: xiii; see OPSI: 61-63). The communicative dimension is performative and refers to a mode of mutual understanding. From a speaker's

¹³⁸ On the continuities between transcendental argument and formal pragmatics, see Power (1993).

¹³⁹ Of all Habermas's programmes, formal pragmatics is most vulnerable to criticism. However, as Finlayson (2005: 40) has argued: 'We should not be tempted to dismiss Habermas's entire philosophy on the grounds that there are errors or misconceptions in his theory of meaning. We should focus, rather, on the question of what insights the pragmatic theory of meaning allows him to bring to social, moral, and political theory'. Habermas describes a shift from the methodological perspective of the Christian Gauss lectures (see OPSI), committed to continuing epistemology by other means, to the substantive social theory of *The Theory of Communicative Action* (see TCA 1: xli, LSS: xiv), in which Habermas admits that he is 'no longer confident that a rigorous transcendental-pragmatic program, claiming to provide ultimate grounds, can be carried out' (TCA 1: 137). It is misleading to suggest that it is only 'in his writings subsequent to *The Theory of Communicative Action*' that Habermas has '[o]n occasion' acknowledged the need for formal pragmatics to be integrated into a social theory of modernity (Cooke 2006: 57).

¹⁴⁰ 'Natural languages can function as their own meta-languages, as Donald Davidson, for example, has also pointed out' (Fultner 2001: xiii).

performative perspective, it is ‘not possible simultaneously to perform and to objectify an illocutionary act’, an act that seeks rationally motivated consensus (OPC: 65). This performative dimension only reveals itself to a hearer who adopts the perspective of a participant entering into the intersubjectively shared lifeworld of a linguistic community (OPC: 217).

Having introduced Habermas’s formal pragmatics from the perspective of the reflexivity of linguistically mediated social practices, I turn next to how Habermas integrates his pragmatic theory of meaning into his mature social theory in *The Theory of Communicative Action*. The pragmatic meaning of speech involves the way speech functions to establish an intersubjective consensus: ‘Reaching understanding is the inherent telos of human speech’ (TCA 1: 287). More specifically, the aim of language is *Verständigung* (understanding/agreement).¹⁴¹ According to what Habermas calls the validity basis of meaning, ‘[s]hared meanings depend on shared reasons’ (Finlayson 2005: 35). ‘Validity claims are internally connected with reasons and grounds’ (TCA 1: 301). Reaching a consensus depends on the intersubjective recognition of validity claims. Any sincere speech act makes three validity claims: truth for a proposition or an existential presupposition, rightness for an action referred to or the norm governing it, and truthfulness for the inner state expressed (TCA 1: 307-308). In being uttered, a sentence is placed in relation to the objective external world, the subjective internal world and the shared intersubjective world (TCA 1: 308; OPC: 49). I explore this further later. The important point for now, though, is that the meaning that the grounding of validity claims has can be explained only in relation to the conditions for discursively redeeming validity claims.

Communicative action provides the link between Habermas’s pragmatic theory of meaning and his social theory. It consists in the threefold relation of an actor reaching understanding, with someone, about something in the world. Habermas argues that communicative action becomes the main basis of social order in modern, secular societies. In communicative action, validity claims are a tacitly accepted basis for coordinating action and remain unquestioned. But hearers always have the choice of adopting a ‘yes’ or ‘no’ position, doing so in light of reasons or grounds that express insight or understanding (TCA 1: 38). Communication can break down when a hearer

¹⁴¹ As Heath (2001: 320 n.42) argues, ‘[f]ar from being an ambiguity, the fact that *Verständigung* suggests both understanding and agreement is precisely the point’, such a connection between truth and understand being one that has also been made by Davidson. ‘The hermeneutic presupposition of rationality exhibits a striking affinity with Davidson’s principle of charity’ (BNR: 65). Habermas now distinguishes between weak and strong senses of communicative action. In the former, actors are oriented towards claims to truth and truthfulness, whereas in the latter they are in addition oriented toward intersubjectively recognised claims to rightness (OPC: 326-29).

demands that the speaker make good the warranty they have implicitly issued to redeem a validity claim in accordance with a speech-act-immanent obligation. In the transition from an action situation to a discourse situation there is a change in attitude as claims to validity are thematized and made problematic. Discourse is the reflexive form of communicative action, communication about communication, in the sense of ‘communication that reflects upon the disrupted consensus in the context of action’ (Finlayson 2005: 41).

Communicative action, which has no other goal beyond the *telos* of mutual understanding inherent to language, is distinguished from strategic action in which action has an independently defined end and speech is employed to get other people to do things as a means to realizing this end. Strategic action is the analogue in relations to persons of instrumental action in manipulating the physical world. Communicative action is prior to strategic and instrumental action since the possibility of a manipulative attitude towards people and things presupposes, and is parasitic upon, a non-manipulative attitude. Habermas’s social ontology seeks to find a place for both action theory and systems theory. The context of communicative action is the lifeworld, while that of strategic action is the system. The lifeworld constitutes the context-forming horizon from within which participants in communication come to an understanding with one another about something. Like Gadamer, Habermas understands the achievement of mutual understanding to have a temporal dynamic and a socio-historical context:¹⁴²

Communicative action can be understood as a circular process in which the actor is two things in one: an initiator who masters situations through actions for which he is accountable and a product of the traditions surrounding him, of groups whose cohesion is based on solidarity to which he belongs, and of processes of socialization in which he is reared (MCCA: 135, see also TCA 2: 135, OPC: 246, 336).

The lifeworld provides the level at which Habermas understands social rationalization to take place. Habermas’s distinction between types of linguistically mediated interaction, associated validity claims and modes of discourse map on to Weber’s distinction between scientific-technical, moral-legal and aesthetic-expressive value spheres (TCA 1: 163-64, 335).

¹⁴² Owen’s (2002) otherwise helpful analysis of Habermas’s idea of social evolution suffers from presenting formal pragmatics as the synchronic framework of Habermas’s critical theory to which the theory of social evolution adds a diachronic framework, whereas I have claimed the two are intertwined.

	Strategic Action	Conversation	Normatively Regulated Action	Dramaturgical Action
World relations	Objective world ('the' world)	Objective world ('the' world)	Intersubjective or social world ('our' world)	Subjective world ('my' world)
Functions of Speech	Influencing one's opposite number	Representation of facts	Establishment of legitimate interpersonal relations	Disclosure of speaker's subjectivity
Action orientations	Success	Reaching understanding	Reaching understanding	Reaching understanding
Validity claims	(Effectiveness)	Truth	Rightness	Truthfulness
Discourse	-	Theoretical	Moral-practical	Aesthetic
Value spheres of lifeworld	-	Scientific-technical	Moral-legal	Aesthetic-expressive

Table 3. Linguistically Mediated Interaction

Source: Based on OPC: 92; TCA 1: 329.

Table 3 summarises the relationship between the elements of Habermas's formal pragmatics, its dimensions of validity, and corresponding types of discourse, and the value spheres of the lifeworld. Habermas's wide frame of reference can sometimes be bewildering, but the distinction between system and lifeworld serves as a useful point of reference. Habermas's conception of communicative reason integrates a Gadamerian hermeneutics of mutual understanding in the lifeworld with a Weberian understanding of the process of social rationalization that takes place for Habermas at the level of the lifeworld as well as at the level of systems. Discourse secures the 'unity of rationality in the multiplicity of value spheres' (TCA 1: 249). This attempt to mediate Gadamer and Weber provides a template for understanding Habermas's attempt to mediate Rawls and Luhmann. Habermas overcomes Luhmann's objections to the model of the top-down normative regulation of social institutions through a conception of principles of right that have their basis in a bottom-up Gadamerian conception of language.

In turning from social theory to normative theory, Habermas can argue that a neo-Hegelian Gadamerian conception of language operating within a neo-Kantian Weberian social theory of modern societies 'frees us from the dilemma of having to choose between Kant and Hegel' (PT: 142). As with the discourse theory of morality so with his discourse theory of law, Habermas picks up the Hegelian aspiration to integrate the abstract universality of justice into the concrete particularism of the common good to redeem this Hegelian aspiration with Kantian means (MCCA: 201).

9.3 Kantianized Hegelianism and Transcendence from Within

It may seem that with the shift from ideology critique to philosophical reconstruction in Habermas's linguistic turn, the critical edge of critical theory has been blunted. In one sense this shift involves adopting a Kantian conception of reflection, rather than the Hegelian conception of reflection discussed in *Knowledge and Human Interests*, a move from phenomenological self-reflection to transcendental reflection (McCarthy 1978: 110). But one must keep in mind the, albeit postmetaphysical, 'Hegelian' context of communicative action out of which 'Kantian' discursive reflection takes shape.¹⁴³ Habermas's move from the transcendental analysis of self-consciousness to the circle of reconstructive sciences is an attempt to close the gap between the transcendental and the empirical, to locate the context-transcending power of reason within social practices (PDM: 297; MCCA: 118). The path of detranscendentalized philosophy is for Habermas a move 'from Kant to Hegel and back again' (TaJ: 175-213). This move

has the effect of relocating the Kantian opposition between the real and the ideal *within* the domain of social practice. Cooperative interaction is seen to be structured around ideas of reason which are neither fully constitutive in the Platonic sense nor merely regulative in the Kantian sense. As idealizing suppositions we cannot avoid making while engaging in processes of mutual understanding, they are actually effective in ways that point beyond the limits of actual situations. As a result, social-practical ideas of reason are both 'immanent' and 'transcendent' to practices constitutive of forms of life (McCarthy, quoted in BNR: 27; see also BFN: 19).

The necessity of the 'must' in the presuppositions that participants *must* make if they are to engage in social practices has a Wittgensteinian-linguistic not Kantian-transcendental character (BNR: 27). For Cohen, rules of regulation as vehicles for realizing justice are regulative of antecedently defined subjective actions. Justice demands that personal behaviour be defensible at the bar of a constitutive Platonic idea of justice. For Kant, practice in the phenomenal world must be brought into line with noumenal regulative ideas. Rawls detranscendentalizes Kant's noumenal world to realize autonomy at the level of rules for the basic structure that are constitutive of a practice of social cooperation. These rules are taken to be regulative of citizens' sense of justice, which has as its object the idea of a well-ordered society. But in seeing society as constituted by structures of meaning, for Habermas rules 'do not serve to regulate a form of behaviour that exists independently of them' (OPSI: 56).

¹⁴³ 'Habermas is the foremost contemporary theorist of Hegelian Geist' (Brandom 2009). See also Baynes (2002a).

This allows Habermas to explain the formation of intersubjectivity without invoking an external observer's perspective. Tradition, in the Gadamerian sense of term, is seen as depending upon ongoing renewals of intersubjectivity taking place through the actions of individuals (LSS: 165).

As agents of communicative action, we are exposed to a transcendence that is integrated in the linguistic conditions of reproduction without being delivered up to it . . . Linguistic intersubjectivity goes beyond the subjects without putting them *in bondage* [*hörig*]. It is not a higher-level subjectivity and therefore, without sacrificing a *transcendence from within*, it can do without the concept of an Absolute (RR: 91).

Habermas has increasingly employed the idea of 'transcendence from within' to describe the simultaneously contextual and context-transcending status of validity claims that he originally sought to represent through the idea of the ideal speech situation.¹⁴⁴ The theory of communicative action reconstructs Hegel's concept of ethical life in postmetaphysical, procedural and intersubjective terms without appealing to a unifying Absolute (PDM: 316). The social world is seen neither as an aggregate of individuals nor as a macrosubject. Rather, it constitutes individuals and is continually reconstituted by them. Habermas finds the conceptual resources for understanding intersubjectivity in this way in George Herbert Mead's model of 'individuation through socialization' (PT: 149-205). No longer do we find, as we do in Rawls, a 'conversion' from a subjective to an intersubjective perspective. Normative grounds are woven into the ground of linguistic social practices.

We are now in a position to understand Habermas's starting point in *BFN*. In gaining access to its object domain through the hermeneutic understanding of meaning, a reconstructive sociology of law must reckon with the tension between facticity and validity (BFN: 19). Acts of position-taking in response to naively accepted validity claims charge the social facts they create with an ideal tension. Validity claims are Janus-faced:

As claims, they transcend any local context; at the same time, they have to be raised here and now and be *de facto* recognized if they are going to bear the agreement of interaction participants that is needed for effective cooperation. The transcendent moment of *universal* validity bursts every provinciality asunder; the obligatory moment of accepted validity claims renders them carriers of a *context-bound* everyday practice (PDM: 322; see also BFN: 21; TaJ: 130).

¹⁴⁴ See OPC: 337-38; TaJ: 84, 221; IO: 7, 41. For analysis, see Rizvi (2007).

Revising his previous position, which Habermas now believes threatened to subordinate politics to morality, *BFN* outlines a branching architectonic. A principle of discourse (D) takes shape in both postconventional morality and positive law, which ‘emerge co-originally from the crumbling edifice of substantial ethical life’ (BFN: 84).¹⁴⁵ The discourse principle is intended to explain the point of view from which norms of action can be impartially justified and is introduced in a manner that aims to be morally neutral (BFN: 108-109, 121, 459). It states:

D: Just those action norms are valid to which all possibly affected persons could agree as participants in rational discourses (BFN: 107; see also IO: 41).

The moral neutrality of the discourse principle does not mean that it lacks *normative* content; indeed it presupposes that practical questions can be judged impartially and decided rationally (BFN: 109). Rather, it means that it is sufficiently abstract that it can assume a different meaning depending upon whether the norm in question is a moral or legal one, thereby issuing in both a moral principle and a principle of democracy (BFN: 105, 233-234).

Habermas argues that the discourse principle reflects ‘those symmetrical relations of recognition built into communicatively structured forms of life in general’ (BFN: 109). This split between morality and legality means that autonomy has a different meaning in these two domains. Autonomy in the moral sphere springs from a single source, whereas in the legal sphere there is a division of roles between authors who make and apply the law and addressees who are subject to valid law (TT: 127). The importance of law, not previously recognised by Habermas, is that it serves as a ‘transmission belt’ (BFN: 54, 354, 448). Law is a ‘medium through which the structures of mutual recognition already familiar from simple interactions and quasi-natural solidarities can be transmitted, in an abstract but binding form, to the complex and increasingly anonymous spheres of a functionally differentiated society’ (BFN: 318). I take up the way the discourse principle takes shape in the form of law in 9.5. However, to further justify my focus on *BFN*, with its concern to synthesise the perspectives of the philosophy of justice and the sociology of law, I want first to argue against theoretical short-cuts that seek to derive a political philosophy of justice directly from Habermas’s discourse theory of truth or discourse theory of morality.

¹⁴⁵ In so doing Habermas further distances himself from Apel. See Apel (2002) and BNR: 77-99.

9.4 Truth, Rightness and Truthfulness

Habermas's 1973 article 'Wahrheitstheorien', and its idea of an ideal speech situation, has 'found a lively resonance' (RC: 273), no doubt encouraged by its chronological, and apparent theoretical, coincidence with Rawls's *TJ*.¹⁴⁶ I will focus on Pettit's analysis of what he calls Habermas's 'consensus theory of justice'. Pettit joins those who believe 'Habermas tries to have his cake and eat it' (Pettit 1980: 182) by dodging the dilemma of choosing between a philosophical defence of an ideal of justice and anti-normative Marxian realism. Habermas's theory 'gives us a criterion of justice, identifying the just social scheme as that which would attract rational consensus [in the ideal speech situation], but it denies that this criterion can be applied with certainty in an imperfect world' (Pettit 1982: 209). The Marxist social critic is left 'gambling on what people would opt for in a rational consensus' (Pettit 1980: 182). Pettit poses the Euthyphro dilemma: does a proposition secure rational consensus because it is true, or is it true because it secures rational consensus (Pettit 1982: 213)? If the former, then truth will be 'non-analytical', while if the latter it will be 'analytical'. Pettit argues that for Habermas truth must be non-analytical as 'Habermas obviously thinks of propositions that attract rational consensus as having a *property* which accounts for that distinction' (1982: 213, emphasis added). By the same token, when it comes to justice, 'norms are not just because they secure a rational consensus; on the contrary, they secure a rational consensus because they are just' (1982: 219).

Pettit recognises his argument assumes a 'biological' model of human needs ascertainable under imperfect circumstances, as opposed to an 'artistic' model under which needs are only fully ascertainable under ideal social circumstances.¹⁴⁷ But in contrasting supposedly objective and purely subjective needs in this way, Habermas's view that needs are intersubjectively constituted is excluded from consideration. Pettit accepts that needs may be something to which each person has privileged access, but argues political philosophy can still propose a 'criterion' of justice. His example is a utilitarian approach that uses an interview technique to determine the outline of a just society (1982: 223-24). One does not have to regard needs as purely subjective to think that modes of data collection bias and shape the articulation of needs. Habermas has always emphasised that public opinion 'must not be confused with survey results' (BFN: 362). Pettit fails sufficiently to note the reflexivity built into discursive justification:

¹⁴⁶ This is despite Habermas's own dissatisfaction and refusal to allow an English translation (Bohman 2008). The article is discussed by Alexy (1989); Fultner (1996).

¹⁴⁷ I follow White (1987: 165-66 n.4) in regarding Pettit's dichotomy as an oversimplification.

‘Even the interpretations of needs in which each individual must be able to recognize what he wants become the object of discursive will-formation’ (LC: 108, quoted in Pettit 1982: 225). There is no ideal observer’s standpoint beyond discourse from which needs can be interpreted.

In taking the concept of truth in a scientific sense as basic and self-evident, Pettit does not engage with Habermas’s linguistic turn. After *Knowledge and Human Interests*, Habermas rejects the idea that epistemology forms the foundation of a critical social theory. It is only in *Truth and Justification* that Habermas has again taken up issues in theoretical rather than practical philosophy (TaJ: 1). Habermas’s early consensus theory of truth is better described, and is later described by Habermas (RC: 273), as a discourse theory of truth (Fultner 1996: 238). At least in retrospect it is evident that Habermas ‘provides a theory not of truth, but of *justification*’ (Fultner 1996: 237). Conceiving of moral truth as analogous to theoretical truth, Pettit assumes that Habermas cannot be committed to a non-realist position that truth is analytic, but approaching matters from the practical perspective of justification, this is just what Habermas assumed. However, Habermas has now given up an epistemic conception of truth (which identifies truth with rational assertability under ideal conditions), arguing that his earlier theory was an ‘overgeneralization of the special case of the validity of *moral* judgements and norms’ (TaJ: 8, emphasis added).

Habermas remarks that ‘particularly in the Anglo-American context, the pragmatic conception of these idealizations [of communicative action] is often misunderstood’ (TaJ: 7). Under the slogan of reconciling Kant with Darwin, Habermas is a pragmatic epistemological realist and a pragmatic moral constructivist. This position, with its consequent differentiation of truth and rightness, is not unproblematic. However, while in a moment I will offer a brief defence, it is worth noting that one could grant Pettit’s claim that it ‘is doubtful whether Habermas has any good reason for not speaking of evaluative truth’ (the truth of moral statements) without accepting Pettit’s inference that truth is thereby a substantive property (Pettit 1982: 210 n.7). Estlund, who follows Pettit’s substantive conception of truth, is similarly inattentive to this possibility. For Estlund, ‘a statement that “*x* is *F*” is true in at least the minimal sense if and only if *x* is indeed *F*’ (Estlund 2008: 25). In arguing against the idea that one can always say that ‘it is *true* that it is right that *x*’, Habermas argues such intuitionist approaches fail because normative statements cannot be tested in the same way as descriptive statements. Moral principles can only be tested from the performative standpoint of a participant rather than the objectivating attitude of an observer (MCCA: 54). Estlund’s minimal conception of truth is reminiscent of Tarski’s Convention T, and Estlund neglects to consider that it could therefore be given a deflationary interpretation. A deflationary

conception of truth provides the opposite perspective to Pettit's and Estlund's from which Joseph Heath (2001) argues against Habermas's distinction between truth and rightness. Rather than equating rightness with truth, Heath presents truth and rightness as alternative elaborations of a single underlying claim of correctness that is internal to language, referring to no substantive property beyond discourse.

That being said, Habermas's recent work provides a strong basis for defending the distinction between truth and rightness, epistemological realism and moral constructivism. Habermas's pragmatism is based on the relationship between justification and learning: reality is something to be coped with not copied. In a circular process, behavioural certainties in the lifeworld are shaken and transformed on the level of argumentation into criticisable validity claims. Validity claims are tested discursively and either vindicated or not, with the result that discursively accepted truths return to the realm of action, producing new behavioural certainties (OPC: 363). Habermas now distinguishes between the truth of a proposition and its rational assertability. We cope with a world presumed to be objective and therefore operate with a notion of unconditional truth, drawing a distinction between believing and knowing. '[T]he process of justification can be guided by a notion of truth that *transcends justification* although it is *always already operatively effective in the realm of action*' (OPC: 372).

While the rightness of moral norms is analogous to the truth of descriptive sentences, it differs in referring to the social rather than the objective world. 'Whereas rational acceptability merely points to the truth of assertoric propositions, it makes a constructive contribution to the validity of moral norms' (IO: 38). Aligning himself with Rawls, Habermas argues that '[m]oral norms, which regulate a reasonable common life among rationally capable subjects in general, are surely not just "discovered" but partly "constructed" at the same time' (BFN: 156). The rightness of moral judgements and norms lacks a justification-transcendent point of reference and so unlike truth, rightness is synonymous with rational justification under ideal conditions.

Facticity consequently means something different in these two domains. Objects in the objective world offer resistance and we therefore presuppose an objective world as a totality of objects rather than of facts (TaJ: 27). 'Facts owe their facticity to their being rooted in a world of objects (about which we state facts) that exist independently of our description of them' (TaJ: 257). By contrast, in the social world we run up against the 'softer stuff' of other minds (TaJ: 42).

[M]oral beliefs do not falter against the resistance of an objective world that all participants suppose to be one and the same. Rather, they falter against the irresolubility of normative dissensus among opposing parties in a shared social world . . . The 'objectivity' of an *other* mind is made of different stuff

than the objectivity of an *unanticipated* reality. The resistance of ‘objective spirit’ can be overcome by moral learning processes that lead the disputing parties to broaden their respective social worlds and to *include* one another in a world they jointly construct in such a way that they can assess their conflicts in the light of shared standards of evaluations and resolve them consensually (TaJ: 256).

The social world is historical; it has a different ontological constitution. But insofar as we test the rightness of moral statements from a universalist point of view,

the reference point of an ideally projected social world of legitimately ordered interpersonal relationships can serve as an equivalent for the absent constraints of an objective world in the course of the presumably rational resolution of moral conflicts of interaction (TaJ: 261).

Pettit’s argument seeks to establish the conclusion that Habermas’s theoretical commitments ‘force him’ to take part in the enterprise of political philosophy (Pettit 1982: 228), in the sense of ‘applying’ a substantive criterion of justice (Pettit 1980: 182). While analytic critics may remain unconvinced by a ‘having one’s cake and eating it’ strategy, Habermas’s answer to the Euthyphro question, like Rawls’s and Scanlon’s, is ‘both’. Truth and justice have a meaning that *transcends* discourse but we can only articulate that meaning *within* discourse. The ideal speech situation always sought to represent the unavoidable presuppositions of procedures of discourse and was never a device for ‘deriving’ substantive principles of justice that would not themselves be discursively justified. Pettit’s argument that there is no tenable middle ground between realist Marxism and purely normative philosophy is, to say the least, unproven.

A similar case of short-circuiting arises when the focus is switched from the ideal speech situation to the principle of universalization (U) proposed by Habermas in his discourse ethics. Interpreters here adopt a Kantian paradigm rather than the Platonic paradigm of Pettit and Estlund, but they too misrepresent Habermas’s view. It has frequently been argued that discourse ethics makes sense only as a theory of democratic legitimacy and not as a theory of morality.¹⁴⁸ This is true to the extent that in detranscendentalizing Kant’s noumenal kingdom of ends, Habermas wants to give to the moral ideal of self-legislation something of its original political meaning as a constructive process extending through time (TaJ: 44). But this is not what critics have in mind. Presupposing that Habermas’s rules of discourse are normative prescriptions, critics dispute whether these rules are applicable beyond the political domain. But discourse ethics ‘does not directly answer the question “What ought I to do?”’

¹⁴⁸ See Wellmer (1991); Heller (1984-85); Jean Cohen (1990); Chambers (1996). See also Murphy (1994).

(Finlayson 2005: 77). It seeks only to reconstruct the impartial moral point of view from which moral questions can be answered. Warning against misconceptions regarding the status of the rules of discourse he borrows from Robert Alexy, Habermas says:

Rules of discourse in Alexy's sense are not *constitutive* of discourses in the sense in which chess rules are constitutive of real chess games. Whereas chess rules *determine* the playing of actual chess games, discourse rules are merely the *form* in which we present the implicitly adopted and intuitively known pragmatic presuppositions of a special type of speech (MCCA: 91).

Habermas never claimed 'the application of Alexy's rules alone can be expected to generate regulative norms' (Kelly 2000: 233). The philosophical task is to make explicit the implicit formal rules of the pragmatic presuppositions that are constitutive of real moral discourse.

In distinguishing between the pragmatic, ethical and moral employments of practical reason, Habermas makes clear that discourse ethics would have been better called a discourse theory of morality. This does not represent a change of heart from the standpoint of a new research programme, but rather develops out of reflections internal to discourse ethics. The problem of the unity of practical reason after it has dissolved into three different forms of argumentation is one that moral theory must 'bequeath . . . unanswered to the philosophy of law' (JA: 17). Moral discourse cannot substitute for law if political self-legislation is to be possible in complex modern societies. If the ideal speech situation is not a device for identifying substantive principles of justice, neither is it an ideal that can be 'approximately realized' (BFN: 322). As a model of 'pure' communicative sociation, it may be employed as a 'foil against which the substratum of *unavoidable* social complexity becomes visible' (BFN: 323), but unavoidable facts in this sense are enabling conditions of, not constraints on, communicative sociation.

In Habermas's genealogical analysis of the cognitive content of morality, published soon after *BFN*, it has become apparent how inextricably bound up with Habermas's theory of modernity is his discourse theory of morality (Finlayson 2000). This genealogy, 'an ideal-typical development that could have taken place under real conditions' (IO: 39), begins from the predicament that members of any moral community are in when making the transition to a modern pluralistic society after a substantive background consensus has been shattered (which Habermas rather unfairly contrasts with an 'original position'). The principle of universalization (U) is inspired by (D) but is initially a proposal arrived at abductively, not a deduction from non-moral premises:

(U) A norm is valid when the foreseeable consequences and side effects of its general observance for the interests and value-orientations of *each individual* could be *jointly* accepted by *all* concerned without coercion (IO: 42; see also BFN: 162).

The ‘discourse-ethical model of justification consists in the derivation of the basic principle (U) from the implicit content of universal presuppositions of argumentation in conjunction with the conception of normative justification in general expressed in (D)’ (IO: 43). This is not viciously circular because the universal presuppositions of argumentation are not normative in a moral sense. This strategy of justifying (U) must be ‘supplemented with genealogical arguments drawing on premises of modernization theory’. ‘With (U) we reassure ourselves in a reflexive manner of a residual normative substance’ – the core of the good within the right – ‘which is preserved in posttraditional societies by the formal features of argumentation and action oriented to reaching a shared understanding’ (IO: 45).

But, harking back to his initial conception of discourse ethics, Habermas also remains committed to Apel’s ‘procedure of establishing universal presuppositions of argumentation by demonstrating performative self-contradictions’ (IO: 45). However, for Habermas such transcendental pragmatic arguments make us aware of the already operating context of communicative action within which discourse is embedded, and this is not historically invariant. Habermas criticises Apel’s aspiration for fundamental grounding (JA: 76-88), seeking instead to elucidate modern moral consciousness within the philosophical discourse of modernity. The discourse theory of morality performs an important task in seeking to explicate the universality of moral validity claims for modern moral agents. Justifying the universality of the need to regulate actions through discourse is not a question for discourse ethics, but rather depends upon modernization theory showing that there is no plausible alternative to this once traditional worldviews become undermined.

9.5 A Discourse Theory of Law and Democracy

I have sought to lend weight to Habermas’s conclusion that ‘[a]n *unmediated* application of discourse ethics (or of an unclarified concept of discourse) to the democratic process leads to muddled analyses’ (BFN: 158). The stage is now set for turning to Habermas’s discourse theory of law. Habermas presents this as an alternative to the social contract tradition, which, he argues, has been unsuccessful in understanding the relationship between morality and law. What he seeks to avoid is the subordination of positive law to natural law in accordance with a ‘Platonic intuition that the legal order imitates the

noumenal order of a ‘kingdom of ends’ and at the same time embodies it in the phenomenal world’ (BFN: 106). For Habermas, by contrast, rights and popular sovereignty are co-original: they emerge equiprimordially. Rawls, and Rawlsians (see Gutmann 2003), have claimed that they too take private and public autonomy to be co-original in this sense, but there are grounds for scepticism regarding this claim. This derives not from Habermas’s narrower point that Rawls is committed to a Kantian conception of natural, or moral, law, but rather from his broader point that this failure results from ‘certain premises rooted in the philosophy of consciousness’ (BFN: 84, 298-99). These failings apply as much to substantive-Rousseauian as they do to Kantian approaches, and Habermas introduces his procedural approach in contrast with both (BFN: 99). I have highlighted how, with its turn to the philosophy of language, Habermas’s reconstructivism provides a new conception of the intersubjective and reflexive features of social practices. On this basis, discourse theory promises to be able to overcome the problems which I have argued beset Rawls’s approach.

As Habermas presents it, both Rousseau and Kant sought, through the notion of autonomy, to show how human rights and popular sovereignty mutually interpret one another. The idea of self-legislation by citizens ‘requires that those subject to law as its addressees can at the same time understand themselves as authors of law’ (BFN: 120). However, neither succeeded, Kant tending towards a liberal prioritisation of human rights and Rousseau towards a republican emphasis on popular sovereignty. Both failings result from the limitations of the paradigm of the philosophy of consciousness (or of the subject, Habermas using the two expressions interchangeably):

Although the premises of the philosophy of the subject allow one to bring reason and will together in the concept of autonomy, one can do so only by ascribing this capacity for self-determination to a subject, be it the transcendental ego of the *Critique of Practical Reason* or the people of the *Social Contract*. If the rational will can take shape only in the individual subject, then the individual’s moral autonomy must reach through the political autonomy of the united will of all in order to secure the private autonomy of each in advance via natural law. If the rational will can take shape only in the macrosubject of a people or nation, then political autonomy must be understood as the self-conscious realization of the ethical substance of a concrete community; and private autonomy is protected from the overpowering force of political autonomy only by the non-discriminatory form of general laws. Both conceptions miss the legitimating force of a discursive process of opinion- and will-formation, in which the illocutionary binding forces of a use of language oriented to mutual understanding serve to bring reason and will together – and lead to convincing positions to which all individuals can agree without coercion (BFN: 103).

Habermas's sympathies, lying as they do with the republican rather than liberal view, justify placing him within a Rousseauian paradigm. Habermas argues that for Rousseau public autonomy does not rest on a foundation of innate moral rights, rather, 'the normative content of human rights enters into the mode of carrying out popular sovereignty' (BFN: 101). Habermas starts from this Rousseauian intuition, but follows Kant in focusing on the medium of positive law. This explains why he refers to his brand of political liberalism as Kantian Republicanism. While Kant subordinates law to morality, Habermas sees autonomy as 'realized in the *medium of law itself*' (BFN: 120). A Rousseauian notion of popular sovereignty is interpreted intersubjectively with reference to the medium of positive law (BFN: 301). The exercise of popular sovereignty refers only to the discursive self-organisation of the legal community and 'does not extend to the whole of society in which the constitutionally organized political system is *embedded*' (BFN: 302). Habermas's focus on positive law derives from the combination of his philosophical pragmatism, with its concern with existing practices, and his sociological understanding of the indispensable function of law in the organisation of complex modern societies: 'Philosophy makes *unnecessary* work for itself when it seeks to demonstrate that it is not simply functionally recommended but also morally required that we organize our common life by means of positive law, and thus that we form legal communities' (BFN: 460).

I turn now to Habermas's reconstruction of the self-understanding of modern legal orders. This begins from the question of what rights citizens must accord one another if they want to legitimately regulate their common life by means of positive law after legitimacy can no longer tenably claim a religious or metaphysical foundation (BFN: 82, 118, 132, 453). Habermas argues that the 'performative *meaning* of this constitution-making practice already contains *in nuce* the entire content of constitutional democracy' (BFN: 453). In a clear allusion to Rawls, Habermas has presented his reconstruction as a two-stage procedure that begins with an 'original condition' in which free and equal persons enter into a constitution-making practice (TT: 124). The participants accept three conditions: First, they jointly resolve to regulate their future association legitimately by means of positive law. Second, they are seen as willing and able to take part in rational discourses, a practice of argumentation with demanding pragmatic presuppositions (TT: 124). This represents the introduction of the discourse principle into the legal form. The initially neutral discourse principle (D) is realized as a principle of democracy, the counterpart in the sphere to law to the principle of universalization (U) in the sphere of morality:

Principle of democracy: only those statutes may claim legitimacy that can meet with the assent (*Zustimmung*) of all citizens in a discursive process of legislation that in turn has been legally constituted (BFN: 110).

Habermas takes a third and final condition to be implicit in the original presentation: a condition of reflexivity. In entering the constitution-making practice, participants are ready to ‘make the meaning of this practice an explicit topic’, to reflect intersubjectively upon the resources that make the performance possible (TT: 124).

In the first stage of the constitution-making procedure, participants identify preconditions that must be met before the work of constitution-making can begin. Because the enterprise is to be realized through the medium of law, conditions must be in place to ensure that every future member of the association has the status of a legal person, a bearer of individual rights. This requires the introduction of three categories of rights, which in this first stage are ‘unsaturated’ or lacking in concrete content:

1. Basic rights that result from the politically autonomous elaboration of the *right to the greatest possible measure of equal individual liberties*.
2. Basic rights that result from the politically autonomous elaboration of the *status of a member* in a voluntary association of consociates under law.
3. Basic rights that result immediately from the *actionability* of rights and from the politically autonomous elaboration of individual *legal protection* (BFN: 122).

There are strong Rawlsian resonances in these three categories of rights, which guarantee the private autonomy of legal subjects who reciprocally recognise one another in their role of addressees of the law. However, Habermas differs from Rawls in his understanding of public autonomy. For Habermas, the constitution-making practice is defined by its ongoing self-determining character. Anticipating that they will be ongoing addressees of the law, participants ‘want to ground an association of citizens who make their own laws’ (TT: 125), a constitution-making practice that has an ongoing self-grounding basis. Participants therefore require a fourth category of rights that enable them to mutually recognise one another not just as addressees but also simultaneously as authors of these rights and of the law in general. This requires:

4. Basic rights to equal opportunities to participate in processes of opinion- and will-formation in which citizens exercise their *political autonomy* and through which they generate legitimate law (BFN: 123).

What is important about this fourth category of rights is that it is

reflexively applied to the constitutional interpretation and the further political development of the basic rights abstractly identified in (1) through (4). For political rights ground the status of free and equal active citizens. This status is self-referential insofar as it enables citizens to change and expand their various rights and duties, or ‘material legal status’, so as to interpret and develop their private and civic autonomy simultaneously (BFN: 123).

All that participants are envisaged to have done in this first stage is to have achieved a philosophical clarification of the nature of the practice in which they are involved. In a second stage, they must ‘step out from behind the veil of empirical ignorance’, assess what needs to be regulated under given historical circumstances and give content to the rights required for this (TT: 126). The self-referential act of institutionalizing civic autonomy is incomplete unless it can stabilise itself in the medium of political power that law tacitly presupposes. In other words, the constitution-making practice must move from being a thought experiment to taking shape in a system of government with law-making as well as law enforcing functions, and with a judiciary. This represents an unfolding of the tension between facticity and validity inside the law: Habermas begins from the idea of legal validity (the tension between the positivity and validity of law), moves through the first stage of the reconstruction of the system of rights (the tension between private and public autonomy), to the second stage of a reconstruction of the principles of the constitutional state in which there is a tension within constitutionally organised political power between coercive state power and its authorisation by legitimate law (BFN: 136). In each case there is a rationality inscribed in law that both points beyond law and is brought to bear on it.¹⁴⁹

At this second stage, Habermas argues that legislation and administrative power depend upon what, following Arendt, he calls communicative power (see also PP: 171-89). In the constitutional state, administrative power depends for its legitimacy on lawmaking, or ‘jurisgenerative’, communicative power. The latter develops in intersubjective procedures of communication in the public sphere in which participants make public use of their reason (BFN: 148). As Habermas summarises it:

A legal order is legitimate to the extent that it equally secures the co-original private and political autonomy of its citizens; at the same time, however, it owes its legitimacy to the forms of communication in which alone this autonomy can express and prove itself (BFN: 409).

¹⁴⁹ ‘The idea of the rule of law sets in motion a spiralling self-application of law, which is supposed to bring the internally unavoidable supposition of political autonomy to bear against the facticity of legally uncontrolled social power that penetrates law *from the outside*’ (BFN: 39).

It is important to note that at this point Habermas is still examining the internal tension between facticity and validity. He is yet to examine the external tension and the consequent problem of how a procedural concept of democracy can be given a sociological translation. I consider this briefly in the next chapter, my main concern here being to explore at a conceptual level the contrast between Habermas's and Rawls's understandings of popular sovereignty. In doing so it will be helpful to review and expand upon the foregoing reconstruction.

In developing a radical democratic interpretation of Kant, Ingeborg Maus has argued that Habermas caricatures the positions of Rousseau and Kant, both of whom also saw the relationship between natural and positive law as co-original and not two-tiered (Maus 2002: 96). In other places Habermas seems more sympathetic to this view (BFN: 472-73), and his contrast between liberal and republican approaches is far more straightforward when put in terms of traditions deriving from Locke and Aristotle (BFN: 497). This framework would have the positive effect of forcing upon Habermas the need for a more nuanced view of Rawls, for he recognises that Rawls's position cannot be equated with Lockean liberalism (BFN: 549 n.10). Maus makes this point in emphasising the circular nature of Habermas's argument: 'The application of the discourse principle to the legal form gives rise "equiprimordially" (on both sides of the circular process) to the principle of democracy, on the one hand, and individual private rights that are identical to the legal code, on the other' (Maus 2002: 97). Or as Habermas puts it:

By means of . . . political autonomy, . . . private autonomy that was first abstractly posited can retroactively assume an elaborated legal shape. Hence the principle of democracy can only appear as the heart of a *system* of rights. The logical genesis of these rights comprises a circular process in which the legal code, or legal form, and the mechanism for producing legitimated law – hence the democratic principle – are *co-originally* constituted (BFN: 122).¹⁵⁰

Putting it in Maus's terms, and in accordance with my presentation of Rawls, when it comes to examining the positions of Rousseau, Kant and Rawls the question is not about which side they fall on a liberal/republican divide. Rather, it is a question of how they understand the point of entry into a circular relationship of rights and popular

¹⁵⁰ See also BFN: 408; Habermas (1998e: 18). '[T]he fact that in Habermas's work subjective liberties are presented as both the result of, and the prerequisite for, democratic lawmaking is not a contradiction but the "whole point of the argument"' (Maus 2002: 94, quoting Klaus Günther). Larmore has moved from criticising Habermas for privileging popular sovereignty over human rights (Larmore 1996: 205-223) to arguing that Habermas's conception of popular sovereignty implicitly presupposes human rights (Larmore 2008: 139-1168). Larmore opposes a circular understanding for not advancing 'the cause of clarity' (Larmore 2008: 163 n.36).

sovereignty, and whether Habermas's discourse theoretical approach offers conceptual advantages in understanding this relationship.

The argument of this chapter has been that it does. Rawls builds in public autonomy, in the form of the fair value of the political liberties, to the liberty principle. In a process of institutionalization, the situation of parties in the original position is to be mirrored in a practice of mutual recognition within social institutions through the formation of an overlapping consensus. Working within a Rousseauian paradigm, Rawls's constructivist Political Liberalism is a Hegelianized Kantianism. Having now presented Habermas's political theory in comparison with Rawls's – a first stage of an original condition of constitution-making and a second stage in which participants step out from behind a veil of empirical ignorance to consider how private autonomy can be stabilised within the institutions of the constitutional state – the differences with Rawls have become clear. While also working within a Rousseauian paradigm, Habermas's reconstructivist Kantian Republicanism is a Kantianized Hegelianism. An original condition of mutual recognition in communicative practices takes shape in the medium of positive law. In both approaches there is a change in perspective, moving from the perspective of a theorist at the first stage to the perspective of participants in the second. In Rawls's constructivism it is the move from the original position to an overlapping consensus; Habermas's reconstructivism first brings the discourse principle into the legal form from outside from the vantage point of the political theorist and then changes perspectives to consider citizens who apply the discourse principle for themselves. In both, in a quasi-dialectical fashion, the order of deduction is not the same as the order of support. While the categories of the rights of private autonomy are first in the order of deduction, the second stage of the construction, or reconstruction, reveals the exercise of public autonomy that gives rights their performative meaning. Rawls and Habermas may share an insight into the constructed character of normative principles, but reconstructivism reconstructs the presuppositions of a self-reconstructing practice. Unlike constructivism, it does not seek to become a moment of that practice and does not therefore assume an initial perspective beyond it. The sense of validity it reconstructs is not absorbed into the facticity of an overlapping consensus in an achievement of reconciliation and collective autonomy. For Habermas, 'citizens make an originary use of a civic autonomy that thereby constitutes itself in a performatively self-referential manner' (BFN: 128). Rights enabling the exercise of popular sovereignty 'do not impose any limitations on what they constitute' (BFN: 128): the constitution-making practice can transcend itself.

Rawls and Habermas both accept the view articulated by Williams, but rejected by Cohen, that ethical thought in relation to the modern world must give up 'the attempt to

find one set of ideas that will represent the demands of ethics in all the spheres to which ethical experience applies' (Williams 2005c: 49). However, both continue to believe that philosophical constructivism or reconstructivism plays an important role in understanding how these spheres are socially unified. This was the importance of the basic structure for Rawls as a division of justificatory labour. But Rawls fails to reconcile the external perspective of the philosopher with the internal perspective of citizens and therefore fails to resolve the problem of self-grounding. I have argued that Habermas's paradigm shift to the philosophy of language allows him to overcome this problem.

The communicative concept of the lifeworld no longer understands society as a whole composed of parts, giving up a concrete conception of justificatory community and notion of 'the people' as an entity (BFN: 80, 301, 150, 185). This is the sense in which Habermas has learned from systems theory. However, like Hegel but for opposite reasons, Luhmann solves the problem of self-grounding too well: all possibility of transcendence is retracted. Habermas pursues a common strategy in *The Philosophical Discourse of Modernity* and *BFN*. His argument is that critics of reason fall into a performative self-contradiction: 'the systems theoretical interpretation of the world as a whole, which leads to a reflexive identity of theory and action, demands the functionalization of the concept of truth, but presupposes a theoretical conception of truth' (Holub 1991: 121). Employing the tools of reason to argue that society has become impervious to reason is self-undermining. With this insight into the performative meaning of communicative rationality, Habermas finds a normative foothold for a process of bootstrapping immanent transcendence. Social rationalization is ambivalent, with the dialectic of systems rationalization setting in motion the conditions for a counteracting dialectic of social learning.

Habermas employs this strategy to argue for the conceptual connection between law and radical democracy. Law cannot legitimate itself, but depends upon communicative power. However, learning from systems theory,

in explaining the democratic process, discourse theory employs a structuralist argument that relieves citizens of the Rousseauian expectation of virtue – the orientation to the common good only needs to be extracted in small increments insofar as practical reason withdraws from the hearts and heads of collective or individual actors into the procedures and forms of communication of political opinion- and will-formation (RSP: 385).

Habermas refers to the individual being unburdened or relieved of the expectation of virtue (BFN: 114). As a functional complement to morality, positive law

relieves the judging and acting person of the considerable cognitive, motivational, and – given the moral division of labor often required to fulfil positive duties – organizational demands of a morality centred on the individual's conscience. Law, as it were, compensates for the functional weakness of a morality that from the observer perspective, frequently delivers cognitively indeterminate and motivationally unstable results (BFN: 452-53).

This is potentially misleading in the sense that a burden once borne by tradition and community *cannot* be borne by individuals. Habermas rejects the idea that the tension between facticity and validity takes the form of a gap between an ideal ethical republic and the reality of routine politics that needs to be bridged (BFN: 281). The effect of an expectation of citizen virtue is to push ‘the democratic process, as it actually proceeds in welfare-state mass democracies, into the pallid light of an instrumentally distorted politics, a “fallen” politics’ (BFN: 277). In modern societies the tension between facticity and validity can only be stabilised by the medium of law, but this means that society can stabilise itself only at the constant risk of instability: ‘the seemingly paradoxical achievement of law is to tame the conflict potential of unleashed individual freedoms through equality-guaranteeing norms, which are compelling only so long as they are recognized as legitimate on the unsteady basis of unleashed communicative freedoms’ (RSP: 452; see also BFN: 130, 145-46). The institutions of democratic government must live off an anarchistic core if they are to guarantee equal liberties for all (BFN: xl). A constitution is a ‘living project that can *endure* only as an ongoing interpretation continually carried forward at all levels of the production of law’ (BFN: 129). In a time of transitions, the purpose of the constitution can only be that of continually realising the system of rights anew in changing circumstances.

10. From Procedure in Substance to Substance in Procedure

In the final section of his ‘Reply to Habermas’, Rawls responds to Habermas’s criticism that justice as fairness is substantive rather than procedural, arguing forcefully that justice is never simply procedural but always relies upon substantive principles. The contrast between procedural and substantive philosophical approaches is a fundamental point of contrast between Rawls and Habermas and an appropriate perspective from which to make a concluding assessment of their respective positions. However, such a comparison is immediately complicated by the fact that this issue is one of the primary examples of how in their dialogue Rawls and Habermas often fail to put their pre-understandings in play.

Most significantly, Rawls does not engage with Habermas’s understanding of justice as ‘nothing material, no determinate “value”, but a dimension of validity’ (AS: 249). For Habermas, Rawls is acting within the competencies of a philosopher when he represents the moral point of view through the device of the original position. However,

as soon as he moves to his two principles, he is speaking as a citizen of the United States . . . There is nothing universal about his particular design for a just society . . . I can’t think of any non-procedural, substantive principles which could be applied to everything at every time (AS: 200-201).

This means that on Habermas’s view, *TJ* is not, or is only partially, a philosophical work (AS: 200). Rawls instead approaches the distinction between procedure and substance in a manner familiar from American legal and democratic theory. Surprisingly, he does not discuss the way in which the original position sought to provide a procedural interpretation of Kant’s categorical imperative, nor the related idea of pure procedural justice. However, on his part Habermas’s presupposition that Rawls is essentially a Kantian leads him to overlook, or minimise, fundamental points of difference. Interpreting Rawls’s original position from the perspective of his own discourse theory of morality, Habermas does not subject to scrutiny how the justification of Rawls’s procedural principles of justice depend upon the substantive background of a shared social world and are intended to be realized within the basic structure that forms that social world.

It is therefore necessary to reframe the debate, moving from a contrast between procedural versus substantive justice to one between procedural and substantive conceptions of practical reason. From this perspective Rawls’s project can be seen to

involve the construction of a procedural interpretation of Kant's categorical imperative within a Hegelian social ontology of a social world formed by the basic structure of society. Rawls's conception of pure procedural justice is inextricably connected with the aspiration to give an ideal form for the basic structure that will maintain background justice. By contrast, Habermas's discourse theories of morality and law reconstruct the idealizing presuppositions of a communicative rationality in which the Hegelian substance of ethical life has been sublimated into Kantian procedures for the public use of reason.

10.1 Procedure and Substance in Justice and in Practical Reason

In responding to Habermas's objection that justice as fairness is substantive rather than procedural, Rawls adopts Hampshire's (1993) distinction between the justice of a procedure and the justice of its outcome (PL: 421). Against Hampshire, Rawls argues that procedural and substantive justice cannot be separated: the justice of a procedure is partly dependent upon the justice of its outcomes. Given that Hampshire accepts that legitimate procedures must be based on minimal substantive values, such as a willingness to let the other be heard, Rawls can convincingly argue that this is not a hard and fast distinction. Rawls goes on to locate the more fundamental issues at stake at the level of a debate between majoritarians and constitutionalists. Democratic pluralists such as Dahl (1989) take a majoritarian position and see basic rights as those that are definitive of democracy and facilitate aspects of the democratic procedure, such as the right to vote, majority rule and freedom of political speech. Constitutionalists, on the other hand, object that this fails to protect rights such as freedom of non-political speech and freedom of religion that are not definitive of the democratic process. On both of these points, Rawls follows J. Cohen's (1994a) critique of Hampshire in which he labels Hampshire a 'democratic pluralist' and connects Hampshire's view with an American constitutional tradition extending through Oliver Wendell Holmes, John Hart Ely and Ackerman, and the normative pluralism of democratic theorists from Harold Laski to Dahl.¹⁵¹

On this basis, Rawls offers a two-pronged response to Habermas. First, he argues that Habermas himself recognises the dependence of democratic procedures on substantive principles when he argues that 'the process of public discussion can be guaranteed to have reasonable outcomes only to the extent that it realizes the conditions

¹⁵¹ Estlund's criticism of a flight from substance in democratic theory echoes Lawrence Tribe's (1985) criticism of, especially Ely's, 'pointless flight from substance' in constitutional theory. Guttmann and Thompson (2002) also assimilate Habermas's concern with 'process' to the terms of legal theory.

of ideal discourse' (PL: 425). Second, insofar as Habermas's view is procedural in focusing on legitimacy rather than justice it is inadequate. The procedural legitimacy of legislative enactments depends upon the justice of the constitution. In the long term, legitimacy is self-undermining, for it allows an undetermined range of injustice that would cause a society to cease to be well-ordered, allowing laws to be passed that would overstep the bounds of justice that they must respect in order to be legitimate (PL: 428-29).

Rawls is certainly right that Habermas's procedural idea of the public use of reason is not free of normative content or purely formal.¹⁵² Indeed, Habermas rejects William Rehg's criticism that his discourse theory of law neglects an aspect of its intrinsic procedural fairness, for example. According to Rehg, quite apart from Habermas's connection of legitimacy to the discourse principle (D), there is intrinsic fairness to procedures such as a coin toss which give each participant an equal chance of winning, or a democratic procedure which gives each participant an equal chance to present their arguments and cast their vote (Rehg 1998: 269). Habermas recognises that the democratic process involves both discourse and decisions, but for Rehg while the two are linked by a 'volitional source prior to all discourses', for Habermas 'the procedures that first establish the internal linkage between discourse and decision draw their legitimating force from the cognitive sources of further discourses' (RSP: 415). Models of procedural democracy, such as Barry's (1979) canonical example of passengers in a railway carriage who must decide whether smoking should be allowed, neglect the importance of a background context in framing the democratic process. Such pared down situations are analogous to a Hobbesian situation where any agreement is better than none, and while the majority method is a reasonable solution, tossing a coin appears equally plausible (Beitz 1989: 76). A more realistic account of democracy quickly points towards the need to justify why the intrinsic features of the democratic procedure have a moral force in connection with principles of fairness and impartiality, and how these conditions can be maintained (Horton 2008: 142).

Habermas does not, then, endorse a procedural model of democracy in Barry's sense. But while for Rawls this entails the endorsement of a constitutionalist position, Habermas's view is not so straightforward. Habermas refers to Ely's argument against judicial activism favourably (BFN: 266), seeing Ely's approach as an example of a process-oriented theory of law that takes a critical-reconstructive approach (LM: 248). Moreover, he takes Dahl's procedural conception of democracy as a starting point for examining the sociological translation of the concept of deliberative politics (BFN: 315-

¹⁵² The serious misunderstandings that result are evident in van Mill (1996); see Alford (1996).

21). However, in both cases this favourable assessment is qualified. While Ely may be right that the formal nature of democratic procedures does not mean that they require supplementation by a substantive theory of rights, as is argued by Lawrence Tribe, he fails to acknowledge that the ‘concept of democratic procedure itself relies on a principle of justice entailing equal respect for all’ (BFN: 266). Ely also fails to develop his constitutional theory in tandem with a theory of democracy. Likewise, Dahl’s positivistic sociological conformation of democratic norms with social reality lacks a conception of deliberative politics developed in terms of ‘rational potentials already at work in society itself’ (BFN: 318). Exploring these aspects of Habermas’s position will allow us to consider whether rejecting the view that the democratic process is purely formal entails endorsing liberal constitutionalism.

Kenneth Baynes is right to warn against reading Habermas’s position through the prism of Anglo-American debates:

Within the context of American discussions, Habermas’s use of the phrase ‘procedural democracy’ could be misleading since it differs from the contrast between procedural and substantive democracy found, for example, in John Ely’s *Democracy and Distrust* or in Brian Barry’s influential essay, ‘Is Democracy Special?’ . . . Habermas’s model is clearly not procedural in this sense since it draws upon the ideals of liberty and equality implicit in the idea of communicative reason (Baynes 2002b: 23).

That Rawls’s presentation of these issues misses the mark is evident in Habermas’s remark that Rawls’s observations on procedural vs. substantive justice ‘do not capture the sense in which I use the expression “procedure” and “procedural rationality” when I assert that a practice of argumentation instituted in a certain way tends to support the assumption that its results are rationally acceptable’ (HR 2: 278 n.18). It will therefore be necessary to look in more detail at Habermas’s conception of procedural rationality before looking at how this conception informs Habermas’s discourse theories of morality and law.

Habermas provides a succinct contrast between his position and Rawls’s in the penultimate paragraph of his initial critique. Habermas says that his procedural moral and legal theory is both more and less modest than Rawls’s as it

focuses exclusively on the procedural aspects of the public use of reason and derives the system of rights from the idea of its legal institutionalization. It can leave more questions open because it entrusts more to the *process* of rational opinion- and will-formation. Philosophy shoulders different theoretical burdens when, as on Rawls’s conception, it claims to elaborate the idea of a just society, while the citizens then use this idea as a platform from which to judge existing arrangements and

policies. By contrast I propose that philosophy limit itself to the clarification of the moral point of view and the procedure of democratic legitimization, to the analysis of the conditions of rational discourses and negotiations. In this more modest role, philosophy need not proceed in a constructive, but only in a *reconstructive* fashion. It leaves substantial questions that must be answered here and now to the more or less enlightened engagement of participants, which does not mean that philosophers may not participate in the public debate, though in the role of intellectuals, not of experts (HR 1: 72).

Rawls's reply acknowledges Habermas's criticism but does not take up Habermas's alternative proposal (PL: 421). (For the time being I leave to one side Habermas's acknowledgement that his position is less modest than Rawls's in its philosophical ambitions). Rawls does not consider the distinction between Habermas's moral and legal theory, and therefore the way in which a common conception of procedural rationality is differently realized within these two domains. This means that Rawls does not consider the way in which Habermas's criticisms are directed against the overall architectonic of justice as fairness. Habermas begins from a criticism of the original position as an operationalization of a Kantian moral point of view, arguing in the first section of his critique that Rawls should keep 'the procedural conception of practical reason free of substantive connotations by developing it in a strictly procedural manner' (HR 1: 57). Here Habermas extends an earlier critique of the way Rawls

builds into the description of the original position those normatively substantive constraints under which the rational egoism of free and equal parties must lead to the choice of correct principles. The fairness of the result is guaranteed by the procedure through which it comes about (LM: 242-43).

This is Rawls's idea of the original position as an instance of pure procedural justice, a concept which curiously Rawls does not take up in his reply.¹⁵³ Nor does Rawls refer to his characterisation of the original position in *TJ* as a 'procedural interpretation of Kant's conception of autonomy and the categorical imperative' (TJ: 256/226).

Habermas's criticisms then turn to Rawls's understanding of how the domain of the political is demarcated. Habermas argues that the normative substance of the domain of the political cannot be identified simply from a moral point of view; in Kantian fashion the 'political' should rather be differentiated in accordance with the criterion of 'legal regulation' (HR 1: 71). This is because 'the normative substance of basic liberal rights is already contained in the indispensable medium for the legal institutionalization

¹⁵³ Indeed, Rawls says he will leave aside the special case of gambling that he had earlier presented as the principal example of pure procedural justice (PL: 421).

of the public use of reason of sovereign citizens' (HR 1: 72). This is the basis of Habermas's contrast between reconstructivism and constructivism that Rawls also fails to address.

Commentators have followed Rawls's lead in failing to address these issues, being more inclined to employ Rawls's concept of pure procedural justice to examine whether Habermas's discourse theories of morality and law are really procedural than to examine the basis of Habermas's commitment to proceduralism in a conception of procedural rationality or examine the place of pure procedural justice in Rawls's approach.¹⁵⁴ The key to understanding the respective positions of Rawls and Habermas is to consider the procedure/substance distinction in the terms of moral philosophy, in which a Kantian procedural approach may be contrasted with a Hegelian substantive approach, rather than the terms of democratic theory:

Democratic theory (D): No appeal to standards beyond those intrinsic to democratic procedures vs. appeal to substantive procedure-independent standards;

Moral theory (M): Formal rational Kantian approach based on procedural reason vs. substantive reasonable Hegelian approach.

As I have argued in analysing Rawls, two dimensions need to be distinguished. First, there is the construction, or reconstruction, of normative grounds. However, second, there is the question of the grounding of these grounds. This requires an account of moral motivation – 'the reason-giving force of moral judgements' (Scanlon 1998: 150) – which is connected to a social ontology and an account of moral psychology and socialization.¹⁵⁵ On the first dimension both Rawls and Habermas are advocates of a procedural ethics (Taylor 1996). On the second dimension, however, Habermas's formalist approach contrasts with Rawls's substantive approach. Rawls's understanding of the intrinsic connection between these two dimensions decisively shapes the overall structure of his approach in a way that Habermas fails to recognise. It is not, as Habermas suggests, that Rawls allows substantive elements to intrude into a procedural approach in a way that could be eliminated without revising the overall structure of the theory.

We may therefore contrast the positions of Estlund, Rawls and J. Cohen, and Habermas in the following terms, in accordance with whether the terms procedure and

¹⁵⁴ See Lafont (2003) on the discourse theory of morality and Rosenfeld (1998a, 1998b) on the discourse theory of law.

¹⁵⁵ This corresponds to Taylor's (1989) distinction between advocacy issues and ontology issues.

substance are to be understood in the terms of democratic theory (D) or of moral theory (M):

Estlund: Substantive_D principles of justice identified by democratic procedures_D;

Rawls and J. Cohen: Substantive_D (= non-procedural_D) principles of justice justified by procedures_M of initial situation and use of public reason within substantive_M institutions;

Habermas: Substantive_D values sublimated into discursive procedures_M for the public use of reason, which form the basis of social order after breakdown of substantive_M worldviews.

We can also see the fundamental distinction between G.A. Cohen and Habermas, despite their superficially similar criticisms of Rawls. Both target Rawls's strategy for relating substantive normative principles and procedures of justification. However, from this shared criticism they draw fundamentally different conclusions: Cohen argues for a purer philosophical analysis focused on identifying substantive principles, while Habermas argues that the competence of philosophy does not extend to the identification of substantive principles but only to reconstructing the procedural presuppositions of appeals to substantive moral principles. In concluding my analysis of Rawls and Habermas, I employ the framework of the relationship between procedure and substance to reiterate my argument for the superiority of Habermas's approach.

10.2 Rawls's Constructivism: Procedure in Substance

In analysing the sense in which Rawls's approach is procedural, it is important to recognise that pure procedural justice is distinct from what Rawls calls formal justice.¹⁵⁶ Rawls defines formal justice as the 'impartial and consistent administration of laws and institutions, whatever their substantive principles' (TJ: 58/51). More broadly, it is a 'kind of minimal justice that is secured when like cases are treated alike' (Nelson 1980: 503). Rawls postpones discussion of whether substantive and formal justice always go together until after he has considered the most reasonable substantive principles of justice. Given that Rawls turns next to presenting his two principles of justice before then considering the topic of procedural justice, different forms of procedural justice may be interpreted as establishing a variety of relationships between formal and

¹⁵⁶ Nelson (1980) emphasises the novelty of Rawls's pure procedural justice. Although Rawls cites Barry's (1965) discussion of procedural justice, Barry does not discuss any examples of what Rawls would classify as pure procedural justice.

substantive principles of justice, the adequacy of formal justice depending upon the substantive background context.

Rawls distinguishes four forms of procedural justice (Table 4).

Form of procedural justice	Procedure-independent criterion?	Procedure guarantees correct outcomes?	Example
Perfect	Yes	Yes	Fair division of a cake
Imperfect	Yes	No	Criminal trial
Pure	No	Yes	Gambling
Quasi-pure	Allowed bounds	Quasi (if outcome within allowed bounds procedure has independent weight)	Legislative enactment of laws and policies

Table 4. Forms of Procedural Justice

Source: Based on Dahl (1989: 165) and Peters (1996: 116).

In the case of perfect procedural justice, there is an independent criterion of justice separate from and prior to the procedure. Assuming that a fair division of a cake is taken to be an equal division, a procedure in which the person who divides the cake is the last to take a slice is guaranteed to lead to the independent standard being met. The plausibility of this example draws upon the existence of an informal social convention about what it means to treat people equally; in this case that when something is to be shared, and there are no reasons to believe anyone has a stronger claim than anyone else, a fair division is an equal division.

Second, in a case of imperfect procedural justice there is again an independent criterion of justice, but this time there is no feasible procedure guaranteed to lead to its being identified. A criminal trial is an example, designed to declare a defendant guilty only if he has committed the offence with which he is charged, but unable to guarantee this will occur.

Third, in pure procedural justice, while there is no independent criterion for the right result, there is a fair procedure that *translates* its fairness to the outcomes provided it has been properly followed. Rawls gives the example of gambling, or, more specifically, the distribution of cash that is the outcome of a series of fair bets. It is significant both that this is a series of *fair* bets – ‘the background circumstances define a fair procedure’ (TJ: 86/75) – and that it is a *series* of fair bets. The procedure has to actually be carried out; we cannot just say that a particular outcome *could* have been produced by a series of fair bets.

Finally, in quasi-pure procedural justice there are allowed bounds for the result and therefore no correct answer for the procedure to identify. However, the fact that the procedure has been carried out lends legitimacy to the outcome. The complexities of this final form are best tackled by considering the context in which it is introduced, namely the legislative enactment of laws and policies. I turn to this after considering the different ways in which pure procedural justice is important in Rawls's theory.

Pure procedural justice applies first to the reasonable constraints on the choice of principles imposed by the veil of ignorance in the original position (TJ: 120, 136/104, 118). It models the rational autonomy of the parties, the fact that they recognise no standpoint external to their own point of view which would constrain them by prior and independent principles of justice (PL: 73). They are constrained only by the formal constraints of the concept of right. While these constraints are Kantian in inspiration, their justification appeals neither to transcendental metaphysics nor conceptual analysis, but rather to the role of principles in 'adjusting the claims that persons make on their institutions and one another' (TJ: 131/113). In calling the original position a procedural interpretation of Kant's conception of autonomy and the categorical imperative, Rawls means that it is set within the framework of a reasonable empirical framework (TJ: not present/226; PL: 285). It takes into account not simply an abstract account of the circumstances of justice and man's place in the world, but the place of persons within a framework of social institutions.

Second, pure procedural justice plays the role, in connection with Rawls's first principle of justice, the liberty principle, of securing the fair value of the political liberties. The idea is to 'incorporate into the basic structure of society an effective political procedure which mirrors in that structure the fair representation of persons achieved by the original position' (PL: 330). It is this, together with the second principle of justice, including the difference principle, which explains why the basic liberties are not merely formal.

We therefore turn, third, to the role of pure procedural justice with respect to distributive shares, which depends upon the fair equality of opportunity secured through the establishment and maintenance of a just basic structure: 'Only against the background of a just basic structure, including a just political constitution and a just arrangement of economic and social institutions, can one say that the requisite just procedure exists' (TJ: 87/76). A just basic structure secures background justice. '[T]he two principles as they work in tandem incorporate an important element of pure procedural justice in the actual determination of distributive shares' (PL: 282).

The connection between the substantive principles of choice in the original position and the fact that these principles are to define an ideal form for the basic

structure of society is of fundamental importance to Rawls's position. A social theory of the basic structure drives the construction of the choice procedure, since '[g]iven the unique features and role of [the basic] structure, the idea of an agreement [in social contract views] must be appropriately transformed if the intent of the Kantian form of the contract doctrine is to be realized' (PL: 288). 'Equality is supported by the general facts of nature and not merely by a procedural rule without substantive force' (TJ: 510/446). It is in this sense that the procedure of the original position takes place within, or, better, against the background of, the substance of a shared social world. Rawls summarises his position as follows:

A purely procedural theory that contained no structural principles for a just social order would be of no use in our world, where the political goal is to eliminate injustice and to guide change toward a fair basic structure . . . In the absence of such an ideal form for background institutions, there is no rational basis for continually adjusting the social process so as to preserve background justice. Thus ideal theory, which defines a perfectly just basic structure, is a necessary complement to nonideal theory without which the desire for change lacks an aim (PL: 285).

Fourth, quasi-pure procedural justice is involved in the ideal legislative procedure at the legislative stage of Rawls's four-stage sequence of application. The outcome of these discussions does not define the right result, which must be within the bounds of justice allowed by the constitution. But if the legislature has enacted policies in ways authorized by a just constitution, the fact of procedural enactment carries weight against those who disagree with the decision and 'cannot convincingly establish their point within the framework of the public conception of justice' (TJ: 362/318).

Finally, the idea of justification through reflective equilibrium may be seen as the most fundamental aspect of pure procedural justice, the formation of an overlapping consensus through the procedure of wide and general reflective equilibrium justifying the use of the procedure of the original position as citizens reflectively endorse a political conception for the basic structure derived from it. I return to this issue after first providing an overview of Habermas's understanding of proceduralism.

10.3 Habermas's Reconstructivism: Substance in Procedure

Habermas's conception of procedural rationality is central to his understanding of the requirements of postmetaphysical thinking in modernity. According to Habermas:

Totalizing thinking that aims at the one and the whole was rendered dubious by a *new type of procedural rationality*, which has asserted itself since the seventeenth century through the empirical methods of the natural sciences, and since the eighteenth century through formalism in moral and legal theory as well as the institutions of the constitutional state. The philosophy of nature and theories of natural law were confronted with a new species of requirements for justification. These requirements shattered *the cognitive privilege of philosophy* (PT: 32).

As a result, both science and morality come to rely solely on the rationality of their procedures: ‘Rationality (*Rationalität*) is reduced to something formal insofar as the rationality (*Vernünftigkeit*) of content evaporates into the validity of results’ (PT: 35). If one translates *Vernünftigkeit* as reasonableness, understood in a Rawlsian sense (Dottori 2003), then we have an important contrast between Rawls’s and Habermas’s positions regarding moral motivation. This does not mean that moral reasoning is purely formal in the sense of scientific instrumental rationality, only that the discourse theory of morality ‘deprives practical reason of all specific normative contents and sublimates it in the form of a procedure for justifying possible normative contents’ (Habermas 1989: 149). This is a formalist ethics insofar as it designates a ‘rule or a procedure that establishes how a morally relevant action conflict can be judged impartially – that is, from a moral point of view’ (Habermas 1990: 228). Kant’s categorical imperative is a prototype, understood not as a maxim of action but as a principle of justification that can be given a procedural form (see Silber 1974).

Like Rawls, Habermas argues that participants in a moral practice have ‘lost their metaphysical guarantees and must so to speak derive their normative orientations from themselves alone’ (IO: 41), drawing on features of a common practice they already share. But in contrast to Rawls’s reliance on a normatively substantive shared social world, Habermas argues that in posttraditional societies, not only can one not expect a substantive consensus on the good, but that these shared features ‘shrink to the fund of formal features of the performatively shared situation of deliberation’ (IO: 41). Joint practical deliberation which justifies moral norms convincing to all participants because of their impartiality becomes a posttraditional equivalent to the ‘traditional, substantive grounding of a normative consensus’. ‘The missing transcendent good can be replaced in an immanent fashion only by appeal to the intrinsic constitution of the practice of deliberation’ (IO: 41). Discourse ethics does ‘have a utopian content, but it does not sketch out a utopia’ (RC: 251). For Habermas, the priority of the right over the good means the uncoupling of the horizontal perspective of the moral regulation of interpersonal relationships from the vertical perspective of ethical reflection on individual and collective life-projects (IO: 28). In our use of practical reason we erect

‘an ideal world not of substantive things but of proper method’ (Mead, quoted in PT: 184). The ‘remnant of the good at the core of the right’ ensures an internal relation between justice and solidarity (IO: 29) as it is a question ‘not so much of two moments that supplement each other as of two aspects of the same thing’ (Habermas 1990: 244).

When it comes to law, Habermas’s position begins from a critique of the social contract tradition and bourgeois formal law, in Weber’s sense, which Habermas takes to depend upon a background theory analogous to the economic assumption of perfect competition leading to an equilibrium in market processes. Habermas criticises Weber for assuming that all forms of procedural rationality are analogous to the formalism of instrumental rationality and ignoring the possibility of ethical formalism. According to Habermas:

Weber does not sufficiently distinguish between structural and substantive – or formal and material aspects [of the social contract]. Only for this reason is he able to mistake ‘nature’ and ‘reason’ for value contents from which formal law was the first to free itself. He falsely identifies the procedural properties of a post-traditional level of justification with substantive values. Therefore, he does not see that the model of the social contract (in a way similar to the categorical imperative) can be understood as proposing a procedure whose rationality is supposed to guarantee the correctness of whatever decisions come about in a procedural manner (LM: 228).

In the balance sheet of modernity, morality has lost its ability to justify substantive ethical positions, but it has gained an integral place within formal rules of social organisation. ‘Morality that is not only complementary to but at the same time ingrained in law is of a procedural nature; it has rid itself of all specific normative contents and has been sublimated into a procedure for the justification of possible normative contents’ (LM: 247).

Procedural democracy cannot be mistaken for a merely formal theory of democracy (BR: 75). But under conditions of postmetaphysical thinking we cannot expect a consensus on substantive issues, only a commitment amongst citizens to legitimately regulate their lives by positive law. This constitution-making practice presupposes law as the medium of binding regulations and the discourse principle as a way of understanding reasonable deliberation and decision making. The restriction to these formal presuppositions is ‘tailored for the specifically modern pluralism of worldviews, cultural forms of life, interest positions, and so forth’ (RSP: 406). Habermas is clear that a constitution-making practice is therefore not free of all normative content. Rawls takes this to imply that it includes substantive elements, pointing to Habermas’s acknowledgement that the procedural legal paradigm retains a dogmatic core in the idea of autonomy (PL: 426). But Habermas argues that

the performative meaning of [the constitution-making practice], which is merely set forth and explicated in constitutional principles and the system of rights, already contains as a doctrinal core the (Rousseauian-Kantian) idea of the self-legislation of voluntarily associated citizens who are both free and equal. This idea is not ‘formal’ in the sense of being ‘value free’. However, it can be fully developed in the course of constitution-making processes that are not based on the previous choice of substantive values, but rather on democratic procedures (RSP: 406).

Rawls has a conventional liberal constitutionalist understanding of the status of majority rule. Because of the ever-present possibility of majority rule leading to unjust outcomes it is a procedure ‘adopted as the most feasible way to realize certain ends antecedently defined by the principles of justice’ (TJ: 361-62/318). For Rawls, no institutional procedure without substantive guidelines for admissible reasons can ‘cancel the maxim “garbage in, garbage out”’ (PL: 431). But Habermas does not understand democracy simply in terms of majority rule, or the idea of exogenous preferences fed into a mechanical procedure. Instead, Habermas argues that constraints on outcomes are constitutive of procedures of discourse. This provides Habermas with resources to argue against Rawls’s concern that his theory is not sufficiently structured to suggest definite results. Majority decisions for Habermas are ‘only caesura in a process of argumentation that has been (temporarily) interrupted under the pressure to decide’ (RSP: 397). They owe their legitimating force to an “imperfect” but “pure” procedural rationality’ (RSP: 397).¹⁵⁷

There remains the objection that it is substantive reasons that are used to judge an outcome correct, reasons independent of a procedure and that render the procedure redundant. Bernard Peters (1996) argues in contradiction of Habermas’s view that ideal discourse is a form of perfect procedural legitimacy, one that under perfect conditions would isolate the correct answer. But this misunderstands Habermas’s view of how practical questions admit of truth. On the one hand, procedures cannot guarantee right answers, but on the other there are no criteria for right answers given prior to argumentation. Form and content interpenetrate: ‘On the one hand, substantive reasons are what convince us that an outcome is right; on the other hand, the soundness of these reasons can be *demonstrated* only in real processes of argumentation’ (RSP: 409). In a superficially paradoxical manner, the substantive content of a moral principle resides in the anticipation that in an ideal procedure it could rebut every possible objection.

¹⁵⁷ Habermas acknowledges these are Rawls’s terms, but not that Rawls never refers to imperfect pure procedural rationality.

Rawls's response to Habermas that public and private autonomy are co-original in justice as fairness at the level of how they are connected as an ideal begs the question (PL: 416-17). It is precisely the idea of such a philosophical construction giving content to the vertical ethical relationship of what Rawls refers to as a 'sovereign democratic people' (PL: 416) that Habermas rejects in favour of reconstructing the idealizing presuppositions of a horizontal relationship of consociates under law and a conception of popular sovereignty as procedure. A political conception of justice may be self-standing, but it grounds itself from the values of a domain of the political secured by shared social institutions.

In the end, Rawls seems to commit himself to the view that the 'people' is an ideal implicit in democratic political culture: that of free and equal persons united together as one legal body, the body politic, which exercises constituent power to make the higher law in such a way that it expresses the political values of public reason, thereby enabling them to realize the (moral) powers that make them free and equal democratic citizens. This conception of the person and the people seems to be the basis for the *substantive conception of democracy* that Rawls sees as implicit . . . in the public political culture of which [the] constitution is an integral part (Freeman 2007b: 210, emphasis added).

The constitution may therefore not be an external imposition on a self-governing people, but in reflecting the original sovereignty of a democratic people, it cannot be experienced by citizens as a project in Habermas's sense. Rawls cannot understand the domain of the political as that of a collective learning process which takes place without the intervention of a substantive philosophical doctrine.

10.4 Procedural Rationality and Religious Pluralism

Rawls's focus on the red herring of procedural versus substantive justice is all the more disappointing when one finds Rawls concluding, near the end of his discussion, that Habermas would not deny that his doctrine is substantive in the same sense as his own, and that Habermas's view must therefore be procedural in a different sense (PL: 277-78). Rawls goes on:

I conjecture . . . that by the terms 'substantive' and 'substantial' [Habermas] means either elements of religious and metaphysical doctrines, or those incorporated in the thought and culture of particular communities and traditions, or possibly both. His main idea, I surmise, is that once the form and structure of the presuppositions of thought, reason, and action – both theoretical and practical – are properly laid out and analyzed by his theory of communicative action, then all the alleged substantial elements of those

religious and metaphysical doctrines and the traditions of communities have been absorbed (or sublimated) into the form and structure of these presuppositions . . . Justice as fairness as a political doctrine wants no part of any such comprehensive account of the form and structural presuppositions of thought and action (PL: 431-32).

Indeed, Rawls acknowledges that justice as fairness is substantive in the sense that it ‘springs from and belongs to the tradition of liberal thought and the larger community of political culture of democratic societies’, and that it therefore fails to be formal and universal in Habermas’s sense, but he argues this is required to respect the limits of a political-not-metaphysical approach (PL: 432). In particular, Rawls takes Habermas to task for what he sees as an occasionally dismissive attitude towards religious and metaphysical views (PL: 376).

How valid is this contention? To be sure, Habermas acknowledges a change in his position from the functionalist description of religion in *The Theory of Communicative Action* (RR: 79). But already by the time of *Postmetaphysical Thinking*, Habermas acknowledged a ‘curious dependence of a philosophy that has forfeited its contact with the extraordinary’:

Philosophy, even in its postmetaphysical form, will be able neither to replace nor to repress religion as long as religious language is the bearer of a semantic content that is inspiring and even indispensable, for this content eludes (for the time being?) the explanatory force of philosophical language and continues to resist translation into reasoning discourses’ (PT: 51).

The way in which Habermas understands postmetaphysical procedural rationality contradicts Rawls’s characterisation. Habermas says:

I use the term ‘postmetaphysical’, not only in a methodological sense that concerns procedures and conceptual means but also in a substantial sense, to describe agnostic positions that make a sharp distinction between belief and knowledge without assuming the validity of a particular religion (as does modern apologetics) or without denying the possible cognitive content of these traditions (as does scientism). Here I want to distinguish between *rationalist* approaches that (in the Hegelian tradition) *subsume* [*aufheben*] the substance of faith into the philosophical concept, from *dialogical* approaches that (following Karl Jaspers) adopt a critical attitude towards religious traditions while at the same time being open to learning from them (BNR: 245).

On this basis, one can turn Rawls’s charge against him. It is a political conception of justice that subsumes the substance of faith in the procedure of an overlapping consensus. To be sure, political liberalism leaves it up to citizens themselves to ‘formulate their own ways of going beyond, or of going deeper, so as to make that

political conception congruent with their comprehensive doctrines' (PL: 377-78). But in accepting a political conception of justice, citizens are accepting principles of social unity informed by, and depending for their full justification upon, a substantive conception of the shared social world formed by the basic structure of society.

Rawls fails to give due credence to Habermas's alternative reconstructive approach. He presents Habermas's view that substantial elements have been sublimated into communicative action as being equivalent to the idea that these substantial elements are 'expressible, or articulated, by' a theory of communicative action (PL: 378). But this fails to take account of the intention of the theory of communicative action to reconstruct an existing reason latent in societies formed by processes of modernization. Habermas seeks to outline how the relationship between faith and knowledge, and discussion between religious and secular citizens, can be understood within a postsecular society, one which must adjust to the fact that 'religious communities continue to exist in a context of ongoing secularization' (FHN: 104). He asks whether the liberal constitutional state is capable of reproducing its normative infrastructure – the motivations on which it depends – from its own secular resources (BNR: 101, 105). Can Rawls's political liberalism meet this challenge? I have argued that it cannot. Rawls's 'division of labor between the political and the metaphysical' (HR 2: 94) is best understood in Hegelian terms. Constitutional democracy for Rawls relies upon a pre-political ethical basis; not to be sure a religious metaphysical backing, but still the background of a social world as the carrier of a shared constitutional tradition, and notwithstanding the fact that political consensus itself relies only on the *form* of comprehensive doctrines shaped by this social world and not upon their particular *content*. Habermas, on the other hand, subscribes to a division of labour between moral theory and the theory of action: 'the conceptual structuring of the contexts of interaction to which questions of political justice refer is not within the province of moral theory' (HR 1: 68).¹⁵⁸

10.5 The Project of Self-grounding and the Exhaustion of Utopian Energies

Why does Habermas fail to recognise and/or fail to criticise what he would surely regard as Rawls's metaphysical communitarianism? I have suggested that the answer lies in the pre-judgements Habermas brings to his debate with Rawls, as a result of which he takes Rawls to be proposing a Kantian moral theory in the mould of his own discourse ethics.

¹⁵⁸ On Habermas on religion, see Lafont (2007); Cooke (2007); Chambers (2007); Boettcher (2009); Lafont (2009); Ferrara (2009) and, in relation to Rawls, Yates (2007).

On the other hand, Habermas does recognise, and critique, the communitarian republicanism of Frank Michelman – one of the leading proponents of applying Rawls's ideas to legal theory in general and constitutionalism in particular – for understanding citizenship in primarily ethical rather than legal terms and seeing politics as the 'reflexive form of substantial ethical life' (BFN: 279, 269).¹⁵⁹ But Michelman is also a major critic of Habermas's view of the constitution as a future-directed project.¹⁶⁰ While sympathetic to the paradigm of deliberative democracy as a way of conceptualising how 'the people' can make the law by which they are governed, Michelman (1997) remains sceptical about whether justification can go all the way down. Circular self-grounding is for him caught in an infinite regress since the power of constitutional authorship cannot itself be legally constituted. A self-grounding constitutional project can find no location in historical time (Michelman 1999).

But Michelman fails to come to terms with how Habermas is proposing a hermeneutical reconstruction of our present historical situation. In reconceiving of popular sovereignty in procedural rather than substantive terms, the regress problem is reposed in future-oriented terms: 'later generations have the task of actualizing the still-untapped normative substance of the system of rights laid down in the original document of the constitution'. This notwithstanding, 'this fallible continuation of the founding event can break out of the circle of a polity's groundless discursive self-constitution only if this process . . . can be understood in the long run as a self-correcting learning process' (TT: 122). When it comes to the contrast between Habermas on the one hand and Rawls and Michelman on the other, one cannot overstate the significance of the differing historical experiences and constitutional traditions of Germany and the United States. But while this raises large and important questions that I cannot address here, I take Habermas to represent a further step in a process of unfolding the potential of the Enlightenment value of autonomy, one more commensurate with 'our' current political situation.

I have endorsed Habermas's view that concepts are best developed in a reciprocal relation with substantive issues. This is not a question of applying the ideal to the real, developing a nonideal theory for 'approximating' ideal discourse (Chambers 1996) or proposing institutional designs for deliberative democracy (Chambers 2003). This is to mistake the level of abstraction at which discourse theory is properly and profitably applied. Rather, I want to look briefly at how the problem of self-grounding takes shape

¹⁵⁹ See Forst's (2002: 100-112) distinction between substantivist and republican communitarianism.

¹⁶⁰ On Habermas and Michelman, see Baynes (1997); Cronin (2006); Olson (2007).

within, and can inform the understanding of, the issues of distributive justice and international justice.

It is not surprising that in diagnosing the exhaustion of utopian energies in a 1984 speech, Habermas should have focused on the crisis of the welfare state (NC: 48-71). Conceptions of distributive justice took shape in the context of a labouring society – this is certainly true of Rawls's conception of property-owning democracy – but maintaining their forward-looking utopian momentum requires coming to terms with the end of this paradigm. This is conspicuously not the case in many debates in the field of so-called luck egalitarianism, but many deliberative democrats have also been loath to make such a shift, as we have seen in discussing J. Cohen. Deliberative democracy has been presented by many as a means of maintaining a traditional socialist agenda for defending substantive principles of distributive justice, with the question 'equality of what?' having to be answered before 'real' deliberation can begin (Knight 1997; Bohman 1997). Habermas's reflexive circular justification of public and private autonomy, on the other hand, provides a framework for a democratic interpretation of the welfare state in terms of enabling a bootstrapping process of political autonomy.¹⁶¹

Finally, there is the question of international justice.¹⁶² I have drawn upon Rawls's account of realistic utopianism in *LP* in order to uncover aspects of Rawls's work that have been there all along but insufficiently recognised. As early as 1958, Rawls can be found describing the universality of justice in terms of the fact that

every *people* may be supposed to have the concept of justice, since in the life of every society there must be at least some relations in which the parties consider themselves to be circumstanced and related as the concept of justice as fairness requires (CP: 71-72, 224 emphasis added).

We could almost call this a *Volksgeist*. Moreover, Rawls's understanding of the moral agency and sense of self-respect of peoples is just one aspect of his social holism. Rawls always used the term 'person' in a 'general way as a subject of claims. In some cases, it means human individuals, but in others it refers to nations, corporations, churches, teams, and so on' (CP: 75). Liberal peoples are united by the 'common sympathies' of a Millian nationalism (LP: 23), while when it comes to decent peoples Rawls's sympathies with Hegel's conception of social unity are apparent in a model of a decent consultation

¹⁶¹ Olson (2006) makes a move in this direction.

¹⁶² On global justice and international relations, see Mertens (2002); Ingram (2003); Moon (2003); Lachapelle (2005); Kreide (2009).

hierarchy based on Hegel's idea of representation by membership of estates (LP: 72-3).¹⁶³

It is not surprising, then, that *LP* shares the strengths and weaknesses of the rest of Rawls's project. In its favour is its attempt to articulate a critical theory (Audard 2007: 273), avoiding realism and relativism on the one hand and utopian cosmopolitanism on the other, and its placing of distributive justice within the context of the republican value of the collective political autonomy of popular sovereignty.¹⁶⁴ To be sure, Rawls fails to come to terms with changing realities, but one can exaggerate the sense in which the Westphalian world has 'vanished' under the pressure of globalization (Buchanan 2000). The real weakness of Rawls's approach lies deeper than this in Rawls's holistic social ontology and conception of social unity achieved through commitment to substantive principles of justice. The Law of Peoples is not positive law, and Rawls's neglect of positive law means he does not concern himself, as Habermas does, with the possible constitutionalisation of international law (DW: 118-47; BNR: 312-53; EFP: 109-31).

We find an even more clear-cut contrast between Rawls and Habermas, and one directly relevant to the problem of self-grounding, in their respective attitudes towards the European Union. Rawls says:

It seems to me that much would be lost if the European union became a federal union like the United States . . . Isn't there a conflict between a large free and open market comprising all of Europe and the individual nation-states, each with its separate political and social institutions, historical memories, and forms and traditions of social policy. Surely these are great value [sic.] to the citizens of these countries and give meaning to their life (Rawls 2003: 16).

Like Rawls, Habermas is attuned to the limitations of philosophy spinning out utopian fantasies. International justice is best pursued through international law, and a postnational constellation through political advocacy. But Habermas sees the 'closure' of collective political life at the level of the nation-state in the era of the American and French revolutions as based on an egalitarian universalism whose claim to validity transcends the borders of the nation-state. In a circular process, 'national consciousness and democratic citizenship reciprocally stabilized each other', but democratic control and political solidarity need not come to a halt at the borders of the nation-state (TT: 100-101). The procedural properties of a public process of discursively structured

¹⁶³ International political theorists are perhaps more receptive to Rawls's Hegelianism. See Sutch (2001); Lin (2008).

¹⁶⁴ The significance of the practice-dependence of Rawls's approach has been rediscovered in the area of global justice. See Freeman (2006c); Sangiovanni (2007); Meckled-Garcia (2008); Ronzoni (2009).

opinion- and will-formation allow democracy to move beyond the need for the pre-political substantial commonalities of a ‘people’ that consolidated the modern the nation-state (PC: 73-74).

In seeking ‘a renewed political closure of an economically unmastered world society’ (PC: 111), the European Union is an exemplary case, and a European constitution could have a catalytic effect. ‘Europe has to apply to itself once again the logic of the circular process in which the democratic state and the nation mutually constituted each other, only this time in a reflexive manner’ (TT: 101). The European project is a marker of hopes for a ‘world domestic policy without a world government’ (PC: 110). In seeking to rejuvenate exhausted utopian energies, the idea of a supranational politics ‘catching up’ with the process of globalization can be conceived as a project, and one that steers a course between progressivist visions of opening and regressive utopias of closure (PC: 54, 88).

Conclusion: Philosophical Proceduralism and Political Radicalism in a Time of Transitions

G.A. Cohen describes Rawls's significance in Hegelian terms. Rawls grasped his age in thought: 'In his work the politics of liberal (in the American sense) democracy and social (in the European sense) democracy rises to consciousness of itself' (RJE: 11). However, the truth requires the addition of another Hegelian thought, that the Owl of Minerva flies only at dusk. For Rawls, socialism remained a live political tradition, the appeal of liberal socialism undiminished after 1989 (LHPP: 323). Philosophy could aspire to render explicit an incipient ideal of democratic community and bolster liberal nationalist self-respect, with civic friendship supported by an institutionally secured sense of justice, and the great value of the integration of individual self-realization into political community in a property-owning democracy trumping the endless pursuit of economic growth in a stationary state, the 'art of living' being more important than the 'art of getting on' (J.S. Mill, quoted in LP: 107 n.33). The background of social solidarity on which Rawls's conception of justice, and its expression of fraternity through the difference principle, depends has an Owl of Minerva quality indeed (Müller 2006: 336). One cannot read Rawls in context and conclude that his position reflects the 'chief values of liberal capitalist civilization' (RJE: 12).

In Rousseauian fashion, Rawls sees distributive equality as constitutive of democratic political autonomy. The assumptions of Rawls's particular approach now appear problematic. Cohen, decrying the greed and fear of liberal capitalist civilization, takes refuge in a utopian defence of socialist equality, seeking, by defending ideals a million miles from the horizon of present political possibility, to preserve faith in this ideal. But placing ideals in an external relation to the procedures of reasoning through which they are identified only threatens to undermine the faith in reason required for those ideas to be realizable. For Habermas, on the other hand, the living legacy of Marxism is a conception of radical democracy. Today socialism can only mean seeking to maintain those utopian energies in an ongoing radical democratic constitutional project that in its openness to the future abjures the dream of a self-grounding reconciling closure.

Leftists 'often find themselves in the role of "conservatives" since they are really trying to preserve values and propagate ideals that have long since been overtaken by a progress they cannot fathom' (Holub 1991: 125). Lying behind Cohen's lament 'Why Not Socialism?' is a defence of small-c conservatism that vests value in things – whether institutions or ideal principles – rather than in relations between persons that stand in a

relation to history (Cohen 2004). Notwithstanding the homely camping trip thought experiment through which he defends the desirability of socialism (WNS), Cohen's commitment to a Platonic paradigm places him in the orbit of postmodern deconstructions of the idea of justice practiced by Jacques Derrida and Christoph Menke, where the deconstructive insight seeks to 'induce us to try even harder to achieve individual justice, conscious of the inevitability of failure' (BNR: 278). The real message of Cohen's critique of Rawls is that justice is an impossible but governing ideal and that it is constructivists who are utopian in holding out the promise of achieving justice. Cohen marches with the vanquished left in a defence of lost causes for which Slavoj Žižek (2008) adopts from Samuel Beckett the motto: 'Try again. Fail again. Fail better'.

Cohen, Rawls and Habermas reconstruct the tradition of classical German philosophy with the tools of analytic and postanalytic philosophy, and I have argued that Habermas's view combines progress in both domains. But one might ask of Habermas 'Can a procedural conception of democracy be radical?' (Chambers 2002). That depends on what one means by radical. It will not be radical in seeking to justify and then apply institutional recipes for democratic deliberation. And it will not contrast utopian norms of distributive equality, seen as the precondition for 'real' democratic deliberation, with existing facts. But do we want political philosophy to be radical in this sense, or to enable us to understand the continued possibilities for a radical reformism, to get to the root of our current political situation?

In reflexively raising the relationship of political theory to our existing social practices we face bracing realities. One is that in increasingly complex societies, academic works of political theory are read by a vanishingly small number of increasingly specialized fellow academics. Political change will not come from seeking to erect Rawlsian egalitarian barricades in the pages of academic journals (Chambers 2006). Rawls's attempt to write as a philosopher-citizen is no longer tenable. Philosophy has become a specialized activity amongst many others and its relationship to practice can only be the indirect one of seeking to redirect the attentions of empirical disciplines to the tension between facticity and validity in existing social practices. A more direct engagement with practice places philosophers in the position of public intellectuals (HR 1: 72).

In this changed guise, philosophy nevertheless maintains its claim to diagnose the self-understanding of modern societies, the interdependence of philosophy and democracy providing the basis for addressing basic normative questions regarding the well-ordered society (TaJ: 290). In seeking to get to the root of our current political situation, Habermas is concerned with nothing less, but also nothing more, than whether

the ideals of the French revolution are still present in particles and fragments of existing reason.

A capitulation of constitutional principles in the face of overwhelming social complexity cannot be ruled out . . . [A] sceptical evaluation of current world conditions is the background for my reflections. This is why my way can be distinguished from purely normative conceptions such as John Rawls's theory of justice, admirable as it is in itself (BR: 132).

But 'moods – and philosophies in a melancholic "mood" – do not justify the defeatist surrender of the radical content of democratic ideals' (BFN: xlii-xliii). In reconfiguring Rawls's economy of philosophical ambition, maintaining utopian optimism of the political will depends upon a dose of realist pessimism of the philosophical intellect. Cohen may not have looked to philosophy for 'ideas that mattered', or for it to 'address itself to the questions agitating the real world' (KMTH: xxi), but we should not do the same.

In common with Rousseau, Kant, Hegel and Marx, for Rawls it is 'not so much the role of moral and political philosophy to tell us how to live our lives or arrange social and political institutions'. Rather, they should 'provide new ways to understand longstanding moral and political traditions and principles, and new ways to argue for (or against) and justify these positions in terms that are amenable to contemporary moral and political consciousness'. Freeman is right that this is 'Rawls's major contribution, and it will be his major legacy' (Freeman 2007a: 459). Rawls's principles of justice are situated within a philosophical system of justification that reflects upon the situatedness of political theory in social practices. In setting out a manifesto for the future of the discipline in 'The Independence of Moral Theory', Rawls concluded by urging:

We must not turn away from this task because much of it may appear to belong to psychology or social theory and not to philosophy. For the fact is that others are not prompted by philosophical inclinations to pursue moral theory; yet this motivation is essential, for without it the inquiry has the wrong focus . . . In this endeavor the aim of those most attracted to a particular view should be not to confute but to perfect (CP: 302).

As Rawls says, quoting Mill, a doctrine 'is not judged at all until it is judged in its best form' (see LHPP: xiii). This means fusing the horizon of its questions with that of our questions, 'understanding what earlier principles require under changing circumstances' (PL: xxxi). We can use Rawls to go beyond Rawls. Going on in a Rawlsian spirit means that faced with an apparent conflict between the political theory of a realistic utopia and recalcitrant social practices, we need to ascend to a higher level of abstraction to see the

roots of the problem, to become more philosophically abstract in order to remain in touch with the real world, to move from realistic utopianism to utopian realism. In a time of transitions, political philosophy must change in order to remain the same.

References

Albert, Hans. 1985. *Treatise on Critical Reason*. Princeton: Princeton University Press.

Alejandro, Roberto. 1993. Rawls's Communitarianism. *Canadian Journal of Philosophy* 23 (1):75-100.

Alexander, Sidney S. 1974. Social Evolution Through Notional Choice. *Quarterly Journal of Economics* 88:597-624.

Alexy, Robert. 1989. *A Theory of Legal Argumentation: The Theory of Rational Discourse as Theory of Legal Justification*. Translated by R. Adler and N. MacCormick. Oxford: Clarendon Press.

Alford, C. Fred. 1996. The Possibility of Rational Outcomes from Democratic Discourse and Procedures: Comment. *Journal of Politics* 58 (3):753-756.

Ameriks, Karl. 2000. *Kant and the Fate of Autonomy*. Cambridge: Cambridge University Press.

Anderson, Perry. 2005. *Spectrum: From Left to Right in the World of Ideas*. London: Verso.

Anscombe, G.E.M. 1958a. On Brute Facts. *Analysis* 18:69-72.

_____. 1958b. Modern Moral Philosophy. *Philosophy* 33:1-19.

Apel, Karl-Otto. 1975. The Problem of Philosophical Fundamental-grounding in Light of a Transcendental Pragmatic of Language. *Man and World* 8 (3):239-275.

_____. 2002. Regarding the Relationship of Morality, Law and Democracy. In *Habermas and Pragmatism*, edited by M. Bookman, M. Aboulafia and C. Kemp. London: Routledge.

Archard, David and Susan Mendus, ed. 2009. *Special Issue of the Journal of Applied Philosophy: Philosophy and Public Policy*.

Aronovitch, Hilliard. 1996. Reflective Equilibrium or Evolving Tradition? *Inquiry* 39 (3/4):399-419.

Ashford, Elizabeth and Tim Mulgan. 2007. Contractualism. In *Stanford Encyclopedia of Philosophy (Winter 2009 Edition)*, edited by E.N. Zalta. <http://plato.stanford.edu/archives/win2009/entries/contractualism>.

Audard, Catherine. 2007. *John Rawls*. Stocksfield: Acumen.

Austin, J.L. 1975. *How to Do Things with Words*. 2nd ed. Cambridge, MA: Harvard University Press.

Avineri, Schlomo. 1969. *The Social and Political Thought of Karl Marx*. Cambridge: Cambridge University Press.

Baker, Lynne Rudder. 2001. Philosophy in Mediis Rebus. *Metaphilosophy* 32 (4):378-394.

Balaguer, Mark. 2003. Can We Know that Platonism is True? *The Philosophical Forum* 34 (3 & 4):459-475.

Barber, Benjamin R. 1989. *The Conquest of Politics: Liberal Philosophy in Democratic Times*. Princeton: Princeton University Press.

Barry, Brian. 1965. *Political Argument*. London: Routledge and Kegan Paul.

_____. 1973. *The Liberal Theory of Justice*. Oxford: Clarendon Press.

_____. 1979. Is Democracy Special? In *Philosophy, Politics, and Society*, edited by P. Laslett. Oxford: Blackwell.

_____. 1989. *Theories of Justice*. Berkeley: University of California Press.

_____. 1995a. *Justice as Impartiality*. Oxford: Oxford University Press.

_____. 1995b. John Rawls and the Search for Stability. *Ethics* 105 (4):874-915.

_____. 2005. *Why Social Justice Matters*. Cambridge: Polity Press.

Baur, Michael. 2002. Reversing Rawls: Criteriology, Contractualism and the Primacy of the Practical. *Philosophy and Social Criticism* 28 (3):251-296.

Baynes, Kenneth. 1992. *The Normative Grounds of Social Criticism: Kant, Rawls, and Habermas*. Albany: State University of New York Press.

_____. 1995. Modernity as Autonomy. *Inquiry* 38 (3):289-303.

_____. 1997. Deliberative Democracy and the Regress Problem: Response to Michelman. *The Modern Schoolman* 74 (4):331-336.

_____. 2002a. Freedom and Recognition in Hegel and Habermas. *Philosophy and Social Criticism* 28 (1):1-17.

_____. 2002b. Deliberative Democracy and the Limits of Liberalism. In *Discourse and Democracy*, edited by R. von Schomberg and K. Baynes. Albany: State University of New York Press.

_____. 2006. Ethos and Institution: On the Site of Distributive Justice. *Journal of Social Philosophy* 37 (2):182-196.

_____. 2007. Freedom as Autonomy. In *The Oxford Handbook of Continental Philosophy*, edited by B. Leiter and M. Rosen. Oxford: Oxford University Press.

Beaney, Michael. 2007. The Analytic Turn in Early Twentieth-century Philosophy. In *The Analytic Turn* edited by M. Beaney. London: Routledge.

_____. 2009. Analysis. In *The Stanford Encyclopedia of Philosophy (Summer 2009 Edition)*, edited by E.N. Zalta. <http://plato.stanford.edu/archives/sum2009/entries/analysis/>.

Bedau, Hugo Adam 1978. Social Justice and Social Institutions. *Midwest Studies in Philosophy* 3:159-75.

Beiner, Ronald. 2009. John Rawls's Genealogy of Liberalism. In *Reflections on Rawls: An Assessment of his Legacy*, edited by S.P. Young. Farnham, Surrey: Ashgate.

Beitz, Charles. 1989. *Political Equality*. Princeton: Princeton University Press.

Benhabib, Seyla. 1982. The Methodological Illusions of Modern Political Theory: The Case of Rawls and Habermas. *Neue Hefte für Philosophie* 21:47-74.

_____. 1994. Deliberative Rationality and Models of Democratic Legitimacy. *Constellations* 1 (1):26-52.

Benson, Peter. 1994. Rawls, Hegel, and Personhood: A Reply to Sibyl Schwarzenbach. *Political Theory* 22 (3):491-500.

Berlin, Isaiah. 2003. The Pursuit of the Ideal. In *The Crooked Timber of Humanity*, edited by H. Hardy. London: Pimlico.

Bertram, Christopher. 1997. Political Justification, Theoretical Complexity, and Democratic Community. *Ethics* 107 (4):563-583.

Bessette, Joseph M. 1981. Deliberative Democracy: The Majority Principle in Republican Government. In *How Democratic is the Constitution?*, edited by R. Goldwin and W. Shambra. Washington DC: American Enterprise Institute.

Blumenberg, Hans. 1985. *The Legitimacy of the Modern Age*. Cambridge, MA: MIT Press.

Boettcher, James W. 2009. Habermas, Religion and the Ethics of Citizenship. *Philosophy and Social Criticism* 35 (1-2):215-238.

Bogen, James. 1995. Internal and External Relations. In *The Oxford Companion to Philosophy*, edited by T. Honderich. Oxford: Oxford University Press.

Bohman, James. 1996. *Public Deliberation: Pluralism, Complexity, and Democracy*. Cambridge, MA: MIT Press.

_____. 1997. Deliberative Democracy and Effective Social Freedom: Capabilities, Resources, and Opportunities. In *Deliberative Democracy: Essays on Reason and Politics*, edited by J. Bohman and W. Rehg. Cambridge, MA: MIT Press.

_____. 1999. Democracy as Inquiry, Inquiry as Democratic: Pragmatism, Social Science, and the Cognitive Division of Labor. *American Journal of Political Science* 43 (2):590-607.

_____. 2005. Critical Theory. In *The Stanford Encyclopedia of Philosophy (Spring 2005 Edition)*, edited by E.N. Zalta. <http://plato.stanford.edu/entries/critical-theory/>.

Bohman, James and William Rehg. 1997. *Deliberative Democracy: Essays on Reason and Politics*. Cambridge, MA: MIT Press.

———. 2008. Jürgen Habermas. In *The Stanford Encyclopedia of Philosophy (Spring 2008 Edition)*, edited by E.N. Zalta. <http://plato.stanford.edu/archives/spr2008/entries/habermas/>.

Boran, Idil. 2005. Rawls and Carnap on Doing Philosophy Without Metaphysics. *Pacific Philosophical Quarterly* 86:459-479.

Bosanquet, Bernard. 2001 [1899]. *The Philosophical Theory of the State*. Kitchener, Ontario: Batoche Books.

Bowie, Andrew. 2003. *Introduction to German Philosophy: From Kant to Habermas*. Cambridge: Polity Press.

Brandom, Robert. 2009. Towards Reconciling Two Heroes: Habermas and Hegel. <http://www.pitt.edu/~brandom/downloads/TRTHHH%2009-5-7%20e.doc>

Brennan, Geoffrey and Philip Pettit. 2005. The Feasibility Issue. In *The Oxford Handbook of Contemporary Philosophy*, edited by F. Jackson and M. Smith. Oxford: Oxford University Press.

Brennan, Geoffrey. 2007. Economics. In *A Companion to Contemporary Political Philosophy*, edited by R. Goodin, P. Pettit and T. Pogge. Oxford: Blackwell Publishing.

Brettschneider, Corey. 2007. *Democratic Rights: The Substance of Self-Government*. Princeton: Princeton University Press.

Brighouse, Harry. 2004. *Justice*. Cambridge: Polity Press.

Brink, David O. 1987. Rawlsian Constructivism in Moral Theory. *Canadian Journal of Philosophy* 17 (1):71-90.

Brudney, Daniel. 1997. Community and Completion. In *Reclaiming the History of Ethics: Essays for John Rawls*, edited by A. Reath, B. Herman and C.M. Korsgaard. New York: Cambridge University Press.

Buchanan, Allen. 2000. Rawls's Law of Peoples: Rules for a Vanished Westphalian World. *Ethics* 110:697-721.

Buchanan, James M. 1972. Rawls on Justice as Fairness. *Public Choice* 13 (1):123-128.

Buchwalter, Andrew. 2001. Political Pluralism in Hegel and Rawls. In *Pluralism and the Pragmatic Turn*, edited by W. Rehg and J. Bohman. Cambridge, MA: MIT Press.

———. 2002. Habermas, Hegel, and the Concept of Law. In *Discourse and Democracy: Essays on Habermas's Between Facts and Norms*, edited by R. von Schomberg and K. Baynes. Albany: State University of New York Press.

Budde, Kerstin. 2009. Constructivism all the Way Down - Can O'Neill Succeed where Rawls Failed? *Contemporary Political Theory* 8:199-223.

Burrow, J. W. 1993. Editor's Introduction. In *The Limits of State Action*, edited by J.W. Burrow. Indianapolis: Liberty Fund.

Callinicos, Alex. 2001. Having Your Cake and Eating It. *Historical Materialism* 9:169-195.

Care, Norman. 1969. Contractualism and Moral Criticism. *The Review of Metaphysics* 23:85-101.

Carens, Joseph. 1981. *Equality, Moral Incentives, and the Market: An Essay in Utopian Politico-Economic Theory*. Chicago: University of Chicago Press.

Cassirer, Ernst. 1945. *Rousseau-Kant-Goethe*. Princeton: Princeton University Press.

———. 1954. *The Question of Jean-Jacques Rousseau*. Translated by P. Gay. Bloomington: Indiana University Press.

Cavell, Stanley. 1990. *Conditions Handsome and Unhandsome*. Chicago: University of Chicago Press.

———. 1999. *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. New ed. Oxford: Oxford University Press.

———. 2005. *Cities of Words*. Cambridge, MA: Harvard University Press.

Chambers, Simone. 1996. *Reasonable Democracy: Jürgen Habermas and the Politics of Discourse*. Ithaca: Cornell University Press.

———. 2002. Can Procedural Democracy be Radical? In *The Political*, edited by D. Ingram. Oxford: Blackwell Publishing.

———. 2003. Deliberative Democratic Theory. *Annual Review of Political Science* 6:307-26.

———. 2006. The Politics of Equality: Rawls on the Barricades. *Perspectives on Politics* 4 (1):81-89.

———. 2007. How Religion Speaks to the Agnostic: Habermas on the Persistent Value of Religion. *Constellations* 14 (2):210-223.

Chapman, John W. 1975. Rawls's Theory of Justice. *American Political Science Review* 69 (2):588-593.

Charney, Evan. 1998. Political Liberalism, Deliberative Democracy, and the Public Sphere. *American Political Science Review* 92 (1):97-110.

Cladis, Mark S. 1994. Wittgenstein, Rawls and Conservatism. *Philosophy and Social Criticism* 20 (1/2):13-37.

Cohen, G.A. 1983. Review of Wood's Karl Marx. *Mind* 92:440-45.

———. 1988. *History, Labour, and Freedom*. Oxford: Oxford University Press.

———. 1989. On the Currency of Egalitarian Justice. *Ethics* 99:906-944.

———. 1993. Equality of What? On Welfare, Goods and Capabilities. In *The Quality of Life*, edited by M.C. Nussbaum and A. Sen. Oxford: Oxford University Press.

———. 1994. Equality as Fact and as Norm: Reflections on the (partial) Demise of Marxism. *Theoria* 83/84:1-11.

———. 1995. *Self-Ownership, Freedom, and Equality*. Cambridge: Cambridge University Press.

———. 1996a. Self-Ownership, History and Socialism: An Interview with G.A. Cohen. *Imprints* 1 (1).

———. 1996b. Reason, Humanity, and the Moral Law. In *The Sources of Normativity*, edited by O. O'Neill. Cambridge: Cambridge University Press.

———. 1997. Back to Socialist Basics. In *Equality*, edited by J. Franklin. London: IPPR.

———. 2000. *If You're an Egalitarian, How Come You're So Rich?* Cambridge, MA: Harvard University Press.

———. 2001 [1978]. *Karl Marx's Theory of History: A Defence*. Expanded ed. Princeton: Princeton University Press.

———. 2004. A Truth in Conservatism: Rescuing Conservatism from the Conservatives. <http://politicaltheoryworkshop.googlepages.com/GACohenConservatism.pdf>.

———. 2008a. *Rescuing Justice and Equality*. Cambridge, MA: Harvard University Press.

———. 2008b. Rescuing Justice from Constructivism and Equality from the Basic Structure Restriction. <http://law.queensu.ca/events/lectureshipsVisitorships/juliusExchangeHarvard2008.pdf>.

———. 2009a. *Why Not Socialism?* Princeton: Princeton University Press.

———. 2009b. Closing Comments. *Rescuing Justice and Equality: Celebrating the Career of G.A. Cohen - Conference at the Centre for the Study of Social Justice (CSSJ)*. Department of Politics and International Relations, University of Oxford. <http://podcasts.ox.ac.uk/>.

Cohen, Joshua, and Joel Rogers. 1995. *Associations and Democracy*. London: Verso.

Cohen, Joshua, and Charles Sabel. 1997. Directly-deliberative Polyarchy. *European Law Journal* 3 (4):313-42.

Cohen, Jean. 1990. Discourse Ethics and Civil Society. In *Universalism vs. Communitarianism*, edited by D. Rasmussen. Cambridge, MA: MIT Press.

Cohen, Joshua. 1982. Review of Cohen, Karl Marx's Theory of History. *Journal of Philosophy* 79 (5).

———. 1989. The Economic Basis of Deliberative Democracy. *Social Philosophy and Policy* 6 (2):25-50.

———. 1991. Democratic Ideals and Substantive Values: Dahl on Democracy. *Journal of Politics* 53 (1):221-225.

———. 1993. Moral Pluralism and Political Consensus. In *The Idea of Democracy*, edited by D. Copp. Cambridge: Cambridge University Press.

———. 1994a. A More Democratic Liberalism. *Michigan Law Review* 92 (6):1503-1546.

———. 1994b. Pluralism and Proceduralism. *Chicago-Kent Law Review* 69 (3):589-618.

———. 1996. Procedure and Substance in Deliberative Democracy. In *Democracy and Difference: Contesting the Boundaries of the Political*, edited by S. Benhabib. Princeton: Princeton University Press.

———. 1997a. The Natural Goodness of Humanity. In *Reclaiming the History of Ethics: Essays for John Rawls*, edited by A. Reath, B. Herman and C.M. Korsgaard. New York: Cambridge University Press.

———. 1997b. Deliberation and Democratic Legitimacy. In *Deliberative Democracy: Essays on Reason and Politics*, edited by J. Bohman and W. Rehg. Cambridge, MA: MIT Press.

———. 1998. Democracy and Liberty. In *Deliberative Democracy*, edited by J. Elster. Cambridge: Cambridge University Press.

———. 1999. Reflections on Habermas on Democracy. *Ratio Juris* 12 (4):385-416.

———. 2002. Taking People as They Are? *Philosophy and Public Affairs* 30 (4):363-386.

———. 2003. For a Democratic Society. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

———. 2004. The Importance of Philosophy: Reflections on John Rawls. *South African Journal of Philosophy* 23 (2):1-6.

Cook, Deborah. 2001. The Talking Cure in Habermas's Republic. *New Left Review* 12:135-151.

Cooke, Maeve. 2002. Five Arguments for Deliberative Democracy. In *Democracy as Public Deliberation*, edited by M. Passerin d'Entrèves. Manchester: Manchester University Press.

———. 2006. *Re-Presenting the Good Society*. Cambridge, MA: MIT Press.

———. 2007. A Secular State for a Postsecular Society? Postmetaphysical Political Theory and the Place of Religion. *Constellations* 14 (2):224 - 238.

Copp, David. 1996. Pluralism and Stability in Liberal Theory. *The Journal of Political Philosophy* 4:191-206.

Cowen, Tyler. 2007. The Importance of Defining the Feasible Set. *Economics and Philosophy* 23:1-14

Cronin, Ciaran. 2006. On the Possibility of a Democratic Constitutional Founding: Habermas and Michelman in Dialogue. *Ratio Juris* 19 (3):343-369.

Dahl, Robert. 1989. *Democracy and its Critics*. New Haven: Yale University Press.

Daniels, Norman. 1989. Equal Liberty and Unequal Worth of Liberty. In *Reading Rawls*, edited by N. Daniels. Stanford: Stanford University Press.

———. 1996. *Justice and Justification: Reflective Equilibrium in Theory and Practice*. Cambridge: Cambridge University Press.

———. 2003. Democratic Equality: Rawls's Complex Egalitarianism. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

Darwall, Stephen L. 1976. A Defense of the Kantian Interpretation. *Ethics* 86 (2):164-70.

———. 2002. Introduction. In *Contractarianism/Contractualism*, edited by S. Darwall. Oxford: Blackwell Publishing.

Davidson, Donald. 2001. On the Very Idea of a Conceptual Scheme. In *Inquiries into Truth and Interpretation*. Oxford: Clarendon Press.

de Francisco, Andrés. 2006. A Republican Interpretation of the Late Rawls. *Journal of Political Philosophy* 14 (3):270-288.

Delaney, C.F. 1975. Rawls on Justice. *The Review of Politics* 37 (1):104-111.

———. 1977. Rawls on Method. *Canadian Journal of Philosophy, Supplementary Volume* 3:153-61.

DeLue, Steven M. 1980. Aristotle, Kant and Rawls on Moral Motivation in a Just Society. *American Political Science Review* 74 (2):385-393.

Dent, Nicholas. 1988. *Rousseau: An Introduction to his Psychological, Social and Political Theory*. Oxford: Basil Blackwell.

———. 2005. *Rousseau*. London: Routledge.

DePaul, M.R. 1986. Reflective Equilibrium and Foundationalism. *American Philosophical Quarterly* 23 (1):59-69.

Dews, Peter. 2001. A Question of Grounding: Reconstruction and Strict Reflexion in Habermas and Apel. In *Critical Theory: Current State and Future Prospects*, edited by P. Uwe Hohendahl and J. Fisher. Oxford and New York: Berghahn Books.

Dietsch, Peter. 2005. Institutions as Motivational Bootstrapping. *USBIG discussion paper*. <http://usbig.net/papers/105-Dietsch-bootstrapping.pdf>.

Doppelt, Gerald. 1990. Beyond Liberalism and Communitarianism: Towards a Critical Theory of Social Justice. In *Universalism vs. Communitarianism*, edited by D. M. Rasmussen. Cambridge, MA: MIT Press.

Dottori, Riccardo, ed. 2003. *Reason and Reasonableness: Vernunft und Vernunftigkeit*. Berlin-Hamburg-Münster: LIT Verlag.

Drebend, Burton. 2003. On Rawls and Political Liberalism. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

Dryzek, John S. 2000. Discursive Democracy vs. Liberal Constitutionalism. In *Democratic Innovation: Deliberation, Representation and Association*, edited by M. Saward. London: Routledge.

———. 2002. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*. Oxford: Oxford University Press.

Dunn, John. 1985. *Rethinking Modern Political Theory*. Cambridge: Cambridge University Press.

———. 1990. Reconceiving the Content and Character of Modern Political Community. In *Interpreting Political Responsibility*. Princeton: Princeton University Press.

Dworkin, Ronald. 1989. The Original Position. In *Reading Rawls: Critical Studies on Rawls' 'A Theory of Justice'*, edited by N. Daniels. Stanford: Stanford University Press.

———. 1996. Objectivity and Truth: You'd Better Believe it. *Philosophy and Public Affairs* 25 (2):87-139.

———. 2001. *Sovereign Virtue*. Cambridge, MA: Harvard University Press.

———. 2002. Ronald Dworkin: Law, Morality, and Equality. In *Constructions of Practical Reason: Interviews on Moral and Political Philosophy*, edited by H. Pauer-Studer. Palo Alto: Stanford University Press.

———. 2003. Equality, Luck and Hierarchy. *Philosophy and Public Affairs* 31 (2):190-98.

Eagleton, Terry. 2007. *The Meaning of Life*. Oxford: Oxford University Press.

Ebertz, Roger P. 1993. Is Reflective Equilibrium a Coherentist Model? *Canadian Journal of Philosophy* 23 (2):193-214.

Ege, Ragip, and Herrade Igersheim. 2008. Rawls with Hegel: The Concept of 'Liberalism of Freedom'. *European Journal of the History of Economic Thought* 15 (1):25-47.

Elster, Jon. 1986. The Market and the Forum. In *The Foundations of Social Choice Theory*, edited by J. Elster and A. Aanund. Cambridge: Cambridge University Press.

_____. 1992. *Local Justice*. New York: Russell Sage Foundation.

_____. 1998. Introduction. In *Deliberative Democracy*, edited by J. Elster. Cambridge: Cambridge University Press.

Estlund, David. 1989. Democracy and the Common Interest: Condorcet and Rousseau Revisited. *American Political Science Review* 83 (4):1317-1322.

_____. 1992. Review of The Normative Grounds of Social Criticism: Kant, Rawls, and Habermas. *Political Theory* 20 (4):694-97.

_____. 1993a. Making Truth Safe for Democracy. In *The Idea of Democracy*, edited by D. Copp, J. Hampton, and J. Roemer. Cambridge: Cambridge University Press.

_____. 1993b. Who's Afraid of Deliberative Democracy? The Strategic/Deliberative Dichotomy in Recent Constitutional Jurisprudence *Texas Law Review* 71:1437-1477.

_____. 1996. The Survival of Egalitarian Justice in John Rawls's *Political Liberalism*. *Journal of Political Philosophy* 4 (1):68-78.

_____. 1997. Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority. In *Deliberative Democracy: Essays on Reason and Politics*, edited by J. Bohman and W. Rehg. Cambridge, MA: MIT Press.

_____. 1998a. The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth. *Ethics* 108 (2):252-275.

_____. 1998b. Liberalism, Equality, and Fraternity in Cohen's Critique of Rawls. *The Journal of Political Philosophy* 6 (1):99-112.

_____. 2001. Introduction. In *Democracy*, edited by D. Estlund. Oxford: Blackwell Publishing.

_____. 2003a. The Audacious Humility of John Rawls. *Dissent* 50 (2).

_____. 2003b. The Democracy/Contractualism Analogy. *Philosophy and Public Affairs* 31 (4):387-412.

_____. 2005. Democratic Theory. In *Oxford Handbook of Contemporary Philosophy*, edited by F. Jackson and M. Smith. Oxford: Oxford University Press.

_____. 2008. *Democratic Authority*. Princeton: Princeton University Press.

Evans, Mark. 2005. Egalitarianism and Merit in a Non-ideal World: The Problem of Two-tier Education. *Politics and Ethics Review* 1:1-21.

Fabre, Cecile. 2007. *Justice in a Changing World*. Cambridge: Polity Press.

Farrelly, Colin. 2000. Incentives and the Natural Duties of Justice. *Politics* 20 (1):19-24.

_____. 2005. Dualism, Incentives and the Demands of Rawlsian Justice. *Canadian Journal of Political Science* 38 (3):675-695.

_____. 2007a. *Justice, Democracy and Reasonable Agreement*. Basingstoke: Palgrave Macmillan.

_____. 2007b. Justice in Ideal Theory: A Refutation. *Political Studies* 55 (4):844-64.

Feinberg, Joel. 1973. Duty and Obligation in the Non-Ideal World. *Journal of Philosophy* 70 (9):263-275.

Ferrara, Alessandro. 2009. The Separation of Religion and Politics in a Post-secular Society. *Philosophy and Social Criticism* 35 (1-2):77-91.

Finlayson, James Gordon. 2000. Modernity and Morality in Habermas's Discourse Ethics. *Inquiry* 43:319-340.

_____. 2005. *Habermas: A Very Short Introduction*. Oxford: Oxford University Press.

_____. 2007. The Habermas-Rawls Dispute Redivivus. *Politics and Ethics Review*. 3 (1):144-162.

Flikschuh, Katrin. 2000. *Kant and Modern Political Philosophy*. Cambridge: Cambridge University Press.

Føllesdal, Dagfinn. 2005. The Emergence of Justification in Ethics. *European Review* 13 (2):169-182.

Forst, Rainer. 2002. *Contexts of Justice: Political Philosophy beyond Liberalism and Communitarianism*. Berkeley: University of California Press.

Franks, Paul. 2007. From Quine to Hegel: Naturalism, Anti-Realism, and Maimon's Question Quid Facti. In *German Idealism: Contemporary Perspectives*, edited by E. Hammer. Abingdon: Routledge.

Frazer, Michael L. 2007. John Rawls: Between Two Enlightenments. *Political Theory* 35 (6):756-80.

Freeman, Samuel. 1991. Contractualism, Moral Motivation, and Practical Reason. *The Journal of Philosophy* 88 (6):281-303.

_____. 1992. Original Meaning, Democratic Interpretation, and the Constitution. *Philosophy and Public Affairs* 21 (1):3-42.

_____. 1994. Utilitarianism, Deontology, and the Priority of Right. *Philosophy and Public Affairs* 19 (2):122-57.

_____. 2000. Deliberative Democracy: A Sympathetic Comment. *Philosophy and Public Affairs* 29:370-418.

_____. 2002. Review of Fred Neuhouser, Foundations of Hegel's Social Theory: Actualizing Freedom. *Ethics* 112 (4):848-854.

_____. 2003. Introduction: John Rawls - An Overview. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

_____. 2006. Moral Contractarianism as a Foundation for Interpersonal Morality. In *Contemporary Debates in Moral Theory*, edited by J. Dreier. Oxford: Blackwell Publishing.

_____. 2007a. *Rawls*. Abingdon: Routledge.

_____. 2007b. *Justice and the Social Contract: Essays on Rawlsian Political Philosophy*. Oxford: Oxford University Press.

_____. 2007c. The Burdens of Public Justification: Constructivism, Contractualism, and Publicity. *Politics, Philosophy and Economics* 6 (1):5-43.

_____. 2009. Original Position. In *The Stanford Encyclopedia of Philosophy (Spring 2009 Edition)*, edited by E.N. Zalta. <http://plato.stanford.edu/archives/spr2009/entries/original-position/>.

Frei, Tamra. 2009. The Redundancy Objection, and Why Scanlon is Not a Contractualist. *Journal of Political Philosophy* 17 (1):47-65.

Fukuyama, Francis. 1989. The End of History? *The National Interest* (Summer).

_____. 1992. *The End of History and the Last Man*. New York: Free Press.

Fultner, Barbara. 1996. The Redemption of Truth: Idealization, Acceptability and Fallibilism in Habermas' Theory of Meaning. *International Journal of Philosophical Studies* 4 (2):233-251.

_____. 2001. Translator's Introduction. In *On the Pragmatics of Social Interaction*. Cambridge: Polity Press.

Gadamer, Hans-Georg. 1976. On the Scope and Function of Hermeneutical Reflection. In *Philosophical Hermeneutics*, edited by D.E. Linge. Berkeley: University of California Press.

_____. 2003 [1960] *Truth and Method*. New York: Continuum.

Galston, William. 1982. Moral Personality and Liberal Theory. *Political Theory* 10 (4):492-519.

_____. 2009. Realism and Moralism in Political Theory: The Legacies of John Rawls. In *Reflections on Rawls: An Assessment of his Legacy*, edited by S.P. Young. Farnham: Ashgate.

Geuss, Raymond. 2008. *Philosophy and Real Politics*. Princeton: Princeton University Press.

Gibbard, Allan. 1991. Constructing Justice. *Philosophy and Public Affairs* 20:264-279.

Giddens, Anthony. 1991. *The Consequences of Modernity*. Cambridge: Polity Press.

_____. 1993. *New Rules of Sociological Method*. 2nd ed. Cambridge: Polity Press.

Gillespie, Michael Allen. 1997. The New Hegel. *Political Theory* 25:584-597.

González, Ana Marta. 2005. John Rawls and the New Kantian Moral Theory. In *The Legacy of John Rawls*, edited by T. Brooks and F. Freyenhagen. London: Continuum.

Goodin, Robert. 1995. Political Ideals and Political Practice. *British Journal of Political Science* 25:37-56.

_____. 2000. Democratic Deliberation Within. *Philosophy and Public Affairs* 29:81-109.

_____. 2001. Managing Scarcity: Toward a More Political Theory of Justice. *Philosophical Issues* 11:202-228.

Goodin, Robert, and Simon J. Niemeyer. 2003. When does Deliberation Begin? Internal Reflection versus Public Discussion in Deliberative Democracy. *Political Studies* 51:627-49.

Goodman, Nelson. 1979. *Fact, Fiction and Forecast*. 4th ed. Cambridge, MA: Harvard University Press.

Gordon, Peter E. 2005. Self-Authorizing Modernity: Problems of Interpretation in the History of German Idealism. *History and Theory* 44:121-137.

Graham, Keith. 1990. Self-Ownership, Communism and Equality. *Proceedings of the Aristotelian Society, Supplementary Volume* 64.

Gray, John. 1991. Contractarian Method, Private Property and the Market Economy. In *Liberalisms*. London: Routledge.

_____. 1994. After the New Liberalism. *Social Research* Fall.

_____. 2000. *Two Faces of Liberalism*. Cambridge: Polity Press.

Grcic, Joseph M. 1983. Kant and Rawls: Contrasting Conceptions of Moral Theory. *Journal of Value Inquiry* 17 (3):235-40.

Griffin, Stephen M. 1993. Political Philosophy versus Political Theory: The Case of Rawls. *Chicago Kent Law Review* 69:691-707.

Guay, Robert. 2006. On Being Constrained by the Facts: Cohen on Facts and Principles. <http://www.columbia.edu/~reg28/cfp.pdf>.

Gunnell, John G. 1996. Review of Theory and Practice: Nomos XXXVII. *American Political Science Review* 90 (2):413-414

Gutmann, Amy. 1989. The Central Role of Rawls's Theory. *Dissent* Summer:338-42.

_____. 2003. Rawls on the Relationship between Liberalism and Democracy. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

Gutmann, Amy, and Dennis Thompson. 2002. Deliberative Democracy Beyond Process. *Journal of Political Philosophy* 10 (2):153-174.

Habermas, Jürgen. 1973. *Theory and Practice*. Translated by J. Viertel. Cambridge: Polity Press.

_____. 1975. *Legitimation Crisis*. Translated by T. McCarthy. Cambridge: Polity Press.

_____. 1976 [1963]. The Analytical Theory of Science and Dialectics. In *The Positivist Dispute in German Sociology*, edited by T.W. Adorno, et al. New York: Harper and Row.

_____. 1976 [1964] A Positivistically Bisected Rationalism. In *The Positivist Dispute in German Sociology*, edited by T.W. Adorno, et al. New York: Harper and Row.

_____. 1979a. *Communication and the Evolution of Society*. Translated by T. McCarthy. Cambridge: Polity Press.

_____. 1979b. History and Evolution. *Telos* 38:5-44.

_____. 1982. A Reply to my Critics. In *Habermas: Critical Debates*, edited by J.B. Thompson, and D. Held. Cambridge, MA: MIT Press.

_____. 1983. *Philosophical-Political Profiles*. Cambridge, MA: MIT Press.

_____. 1984. *The Theory of Communicative Action: Vol. I, Reason and the Rationalization of Society*. Translated by T. McCarthy. Cambridge: Polity Press.

_____. 1986. Law and Morality. *The Tanner Lectures on Human Value*. <http://www.tannerlectures.utah.edu/lectures/documents/habermas88.pdf>.

———. 1987a. *Knowledge and Human Interests*. Translated by J.J. Shapiro. Cambridge: Polity Press.

———. 1987b. *The Philosophical Discourse of Modernity*. Translated by F. Lawrence. Cambridge: Polity Press.

———. 1987c. *The Theory of Communicative Action: Vol.2, Lifeworld and System: A Critique of Functional Reason*. Translated by T. McCarthy. Cambridge: Polity Press.

———. 1988. *On the Logic of the Social Sciences*. Translated by S.W. Nicholsen and J.A. Stark. Cambridge: Polity Press.

———. 1989. Towards a Communication-concept of Rational Collective Will-formation. A Thought-experiment. *Ratio Juris* 2 (2):114-54.

———. 1989 [1962]. *The Structural Transformation of the Public Sphere*. Translated by T. Burger, and F. Lawrence. Cambridge: Polity Press.

———. 1990. Justice and Solidarity: On the Discussion concerning Stage 6. In *The Moral Domain: Essays in the Ongoing Discussion Between Philosophy and the Social Sciences*, edited by T.E. Wren. Cambridge, MA: MIT Press.

———. 1990a. What Does Socialism Mean Today? The Rectifying Revolution and the Need for New Thinking on the Left. *New Left Review* 183.

———. 1990b. *Moral Consciousness and Communicative Action*. Translated by C. Lenhardt and S.W. Nicholsen. Cambridge: Polity Press.

———. 1990c. *The New Conservatism*. Translated by S.W. Nicholsen. Cambridge: Polity Press.

———. 1992a. *Autonomy and Solidarity*. Edited by P. Dews. Revised and enlarged ed. London: Verso.

———. 1992b. *Postmetaphysical Thinking*. Translated by W.M. Hohengarten. Cambridge: Polity Press.

———. 1993. *Justification and Application: Remarks on Discourse Ethics*. Translated by C. Cronin. Cambridge: Polity Press.

———. 1994. *The Past as Future*. Translated by M. Pensky. Cambridge: Polity Press.

———. 1996. *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. Translated by W. Rehg. Cambridge: Polity Press.

———. 1997. *A Berlin Republic: Writings on Germany*. Translated by S. Rendall. Cambridge: Polity Press.

———. 1998a. *On the Pragmatics of Communication*. Translated by B. Fultner. Cambridge: Polity Press.

———. 1998b. *The Inclusion of the Other: Studies in Political Theory*. Edited by C. Cronin and P. DeGreiff. Cambridge: Polity Press.

———. 1998c. Reply to Symposium Participants, Benjamin N. Cardozo School of Law. In *Habermas on Law and Democracy: Critical Exchanges*, edited by M. Rosenfeld. Berkeley and Los Angeles: University of California Press.

———. 2001. *The Postnational Constellation*. Translated by M. Pensky. Cambridge: Polity Press.

———. 2002a. *On the Pragmatics of Social Interaction: Preliminary Studies in the Theory of Communicative Action*. Translated by B. Fultner. Cambridge: Polity Press.

———. 2002b. *Religion and Rationality: Essays on Reason, God, and Modernity*. Edited by E. Mendieta. Cambridge: Polity Press.,

———. 2003a. *The Future of Human Nature*. Translated by W. Rehg, M. Pensky, and H. Beister. Cambridge: Polity Press.

———. 2003b. *Truth and Justification*. Translated by B. Fultner. Cambridge: Polity Press.

———. 2006a. *The Divided West*. Translated by C. Cronin. Cambridge: Polity Press.

———. 2006b. *Time of Transitions*. Edited by C. Cronin and M. Pensky. Cambridge: Polity Press.

———. 2008. *Between Naturalism and Religion: Philosophical Essays*. Translated by C. Cronin. Cambridge: Polity Press.

———. 2009. *Europe: The Faltering Project*. Translated by C. Cronin. Cambridge: Polity Press.

Hampshire, Stuart. 1993. Liberalism: The New Twist. *New York Review of Books* 40 (14).

Hampton, Jean. 1980. Contracts and Choices: Does Rawls Have a Social Contract Theory? *Journal of Philosophy* 77 (6):315-338.

———. 1989. Should Political Philosophy be Done Without Metaphysics? *Ethics* 99 (4):791-814.

———. 1994. The Common Faith of Liberalism. *Pacific Philosophical Quarterly* 75:186-216.

Hanley, Ryan Patrick. 2004. Political Science and Political Understanding: Isaiah Berlin on the Nature of Political Inquiry. *American Political Science Review* 98 (2):327-339.

Hardimon, Michael O. 1994. *Hegel's Social Philosophy: The Project of Reconciliation*. Cambridge: Cambridge University Press.

Hare, R.M. 1989. Rawls' Theory of Justice. In *Reading Rawls*, edited by N. Daniels. Stanford: Stanford University Press.

Hart, H.L.A. 1955. Are There Any Natural Rights? *Philosophical Review* 64.

———. 1997 [1961]. *The Concept of Law*. 2nd ed. Oxford: Oxford University Press.

Hartmann, Klaus. 1972. Hegel: A Non-Metaphysical View. In *Hegel: A Collection of Critical Essays*, edited by A. MacIntyre. Garden City, NY: Anchor Books.

Haslett, D.W. 1987. What is Wrong with Reflective Equilibrium? *Philosophical Quarterly* 37:305-311.

Heath, Joseph. 2001. *Communicative Action and Rational Choice*. Cambridge, MA: MIT Press.

———. 2009. Habermas and Analytical Marxism. *Philosophy and Social Criticism*, 35 (8):891-919.

Hegel, G.W.F. 1991 [1820]. *Elements of the Philosophy of Right*. Edited by A. Wood. Cambridge: Cambridge University Press.

Heller, Agnes. 1984-85. The Discourse Ethics of Habermas: Critique and Appraisal. *Thesis Eleven* 10/11:5-17.

Henrich, Dieter. 1999. What is Metaphysics - What is Modernity? Twelve Theses against Jürgen Habermas. In *Habermas: A Critical Reader*, edited by P. Dews. Oxford: Blackwell.

Hill, Thomas E. 1989. Kantian Constructivism in Ethics. *Ethics* 99 (4):752-770.

Höffe, Ottfried. 1984. Is Rawls' Theory of Justice Really Kantian? *Ratio* 26:103-24.

Holbo, John. 2008. Rawls and 'Liberalism'. *Crooked Timber*. 7 February. <http://crookedtimber.org/2008/02/07/rawls-and-liberalism/>.

Holub, Robert C. 1991. *Jürgen Habermas: Critic in the Public Sphere*. London: Routledge.

Honig, Bonnie. 1993. *Political Theory and the Displacement of Politics*. Cornell: Cornell University Press.

———. 2007. Between Decision and Deliberation: Political Paradox in Democratic Theory. *American Political Science Review* 101 (1):1-17.

Honneth, Axel. 2007. *Disrespect: The Normative Foundations of Critical Theory*. Cambridge: Polity Press.

Hooker, Brad. 2000. *Ideal Code, Real World: A Rule-consequentialist Theory of Morality*. Oxford: Oxford University Press.

Horkheimer, Max. 1972 [1937]. Traditional and Critical Theory. In *Critical Theory: Selected Essays*. New York: Continuum.

Horkheimer, Max and Theodor W. Adorno. 1997 [1944]. *Dialectic of Enlightenment*. London: Verso.

Horton, John. 2008. Proceduralism as Thin Universalism: Stuart Hampshire's 'Procedural Justice'. In *Principles and Political Order: The Challenge of Diversity*, edited by B. Haddock, P. Roberts and P. Such. London: Routledge.

Houlgate, Stephen. 2001. Hegel, Rawls and the Rational State. In *Beyond Liberalism and Communitarianism: Studies in Hegel's Philosophy of Right*, edited by R. R. Williams. Albany: State University of New York Press.

Howard, Dick. 1996. Law and Political Culture. *Cardozo Law Review* 17 (4-5):1391-1430.

Hoy, Joyce Beck. 1981. Hegel's Critique of Rawls. *Clio* 10 (4):407-422.

Huemer, Michael. 1996. Rawls's Problem of Stability. *Social Theory and Practice* 22 (3):375-395.

Humboldt, Wilhelm von. 1993. *The Limits of State Action*. Indianapolis: Liberty Fund.

Hylton, Peter. 1990. *Russell, Idealism, and the Emergence of Analytical Philosophy*. Oxford: Clarendon Press.

———. 1993. Hegel and Analytic Philosophy. In *The Cambridge Companion to Hegel*, edited by F.C. Beiser. Cambridge: Cambridge University Press.

Ingram, David. 2003. Between Political Liberalism and Postnational Cosmopolitanism: Toward an Alternative Theory of Human Rights. *Political Theory* 31 (3):359-391.

Inwood, Michael. 1992. *A Hegel Dictionary*. Oxford: Blackwell.

Ivison, Duncan. 2002. *Postcolonial Liberalism*. Cambridge: Cambridge University Press.

Jackson, Ben. 2005. Revisionism Reconsidered: Property-Owning Democracy and Egalitarian Strategy in Post-War Britain. *Twentieth Century British History* 16 (4):416-40.

James, Aaron. 2005. Constructing Justice for Existing Practices: Rawls and the Status Quo. *Philosophy and Public Affairs* 33 (3):281-316.

Jay, Martin. 1996. *The Dialectical Imagination*. Paperback ed. Berkeley: University of California Press.

Jensen, Mark. 2009. The Limits of Practical Possibility. *Journal of Political Philosophy* 17 (2):168-84.

Johnson, Oliver A. 1974. The Kantian Interpretation. *Ethics* 85 (1):58-66.

———. 1977. Autonomy in Kant and Rawls: A Reply. *Ethics* 87 (3):251-54.

Johnson, Pauline. 2005. Habermas: A Reasonable Utopian? *Critical Horizons* 6 (1):101-18.

Josefson, Jim and J. Bach. 1997. A Critique of Rawls's Hermeneutics as Translation. *Philosophy and Social Criticism* 23 (1):99-124.

Julius, A.J. 2003. Basic Structure and the Value of Equality. *Philosophy and Public Affairs* 31 (4):321-355.

Kain, Patrick. 1999. Kant's Moral Constructivism and his Conception of Legislation. *The Paideia Project Online: Proceedings of the Twentieth World Congress of Philosophy*. <http://www.bu.edu/wcp/Papers/TEth/TEthKain.htm>.

———. 2004. Self-legislation in Kant's Moral Philosophy. *Archiv für Geschichte der Philosophie* 86 (3):257-306.

Kamm, Frances M. 1996. High Theory, Low Theory, and the Demands of Morality. In *Theory and Practice: Nomos XXXVII*, edited by I. Shapiro and J.W. DeCew. New York: New York University Press.

Kant, Immanuel. 1991 [1785]. *Groundwork of the Metaphysic of Morals*. Translated by H. J. Paton. London: Routledge.

———. 1991 [1784]. What is Enlightenment? In *Kant: Political Writings*, edited by H. S. Reiss. Cambridge: Cambridge University Press.

———. 1991 [1793]. On the Common Saying: 'This May be True in Theory, but it does not Apply in Practice'. In *Kant: Political Writings*, edited by H. S. Reiss. Cambridge: Cambridge University Press.

Karp, David Jason. 2009. Facts and Values in Politics and Searle's *Construction of Social Reality*. *Contemporary Political Theory* 8 (2):152-175.

Katz, Jerry. 1998. *Realistic Rationalism*. Cambridge, MA: MIT Press.

Kaufman, Alexander. 2004. The Myth of the Patterned Principle: Rawls, Nozick and Entitlements. *Polity* 36 (4):559-578.

———. 2006. Rawls's Practical Conception of Justice: Opinion, Tradition and Objectivity in Political Liberalism. *Journal of Moral Philosophy* 3:23-43.

Kelleher, Paul. 2008. How to Read *Political Liberalism*.
http://paulkelleher.net/files/Kelleher_How%20to%20Read%20PL.pdf.

Kelly, Christopher. 1987. 'To Persuade without Convincing': The Language of Rousseau's Legislator. *American Journal of Political Science* 31 (2):321-335.

Kelly, Erin. 2000. Habermas on Moral Justification. *Social Theory and Practice* 26 (2):223-49.

Kelly, Paul. 2000. Political Theory in Retreat? Contemporary Political Theory and the Historical Order. In *Political Theory in Transition*, edited by N. O'Sullivan. London: Routledge.

———. 2006. The Social Theory of Anti-Liberalism. *Critical Review of International Social and Political Philosophy* 9 (2):137-154.

Kelly, Terrence. 2001. Sociological Not Political: Rawls and the Reconstructive Social Sciences. *Philosophy of the Social Sciences* 31 (1):3-19.

King, Desmond. 2008. Symposium: Contract and Domination by Carole Pateman and Charles W. Mills. *Journal of Political Ideologies* 13 (3):227-262.

Klosko, George. 1997. Political Constructivism in Rawls's Political Liberalism. *The American Political Science Review* 91 (3):635-646.

Knight, Jack and James Johnson. 1997. What Sort of Equality Does Deliberative Democracy Require? In *Deliberative Democracy: Essays on Reason and Politics*, edited by J. Bohman and W. Rehg. Cambridge, MA: MIT Press.

Knowles, Dudley. 2002. *Routledge Philosophy Guidebook to Hegel and the Philosophy of Right*. London: Routledge.

Kohlberg, Lawrence. 1981. Justice as Reversibility: The Claim to Moral Adequacy of a Highest Stage of Moral Judgment. In *The Philosophy of Moral Development: Moral Stages and the Idea of Justice*. San Francisco: Harper & Row.

Korsgaard, Christine. 1996a. *The Sources of Normativity*. Cambridge: Cambridge University Press.

———. 1996b. The Reasons we can Share: An Attack on the Distinction between Agent-relative and Agent-neutral Values. In *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.

———. 1996c. Two Distinctions in Goodness. In *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.

———. 2008. Realism and Constructivism in Twentieth Century Moral Philosophy. In *The Constitution of Agency*. Oxford: Oxford University Press.

Kortian, Garbis. 1980. *Metacritique: The Philosophical Argument of Jürgen Habermas*. Cambridge: Cambridge University Press.

Krasnoff, Larry. 1998. Consensus, Stability, and Normativity in Rawl's Political Liberalism. *The Journal of Philosophy* 95 (6):269-292.

———. 1999. How Kantian is Constructivism? *Kant-Studien* 90 (4):385-409.

Kreide, Regina. 2009. Preventing Military Humanitarian Intervention? John Rawls and Jürgen Habermas on a Just Global Order. *German Law Journal* 1:93-113.

Krouse, Richard and M. McPherson. 1988. Capitalism, 'Property-Owning Democracy,' and the Welfare State. In *Democracy and the Welfare State*, edited by A. Gutmann. Princeton: Princeton University Press.

Kuhn, Thomas S. 1970. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press.

Kukathas, Chandran and P. Pettit. 1990. *Rawls: A Theory of Justice and its Critics*. Cambridge: Polity Press.

Kymlicka, Will. 1990. *Contemporary Political Philosophy: An Introduction*. Oxford: Oxford University Press.

Lachapelle, Erick. 2005. Morality, Ethics and Globalization: Lessons from Kant, Hegel, Rawls, and Habermas. *Perspectives on Global Development and Technology* 4 (3-4):603-644.

Lackey, Douglas. 2003. Why Not Platonism? In Ethics, Anyway. *The Philosophical Forum* 34 (3 & 4):391-399.

Laden, Anthony Simon. 2001. *Reasonably Radical: Deliberative Liberalism and the Politics of Identity*. Ithaca: Cornell University Press.

———. 2003. The House That Jack Built: Thirty Years of Reading Rawls. *Ethics* 113 (1):367-390.

———. 2004. Taking the Distinction between Persons Seriously. *Journal of Moral Philosophy* 1 (3):277-292.

———. 2006. Republican Moments in Political Liberalism. *Revue Internationale de Philosophie* 60:339-366.

———. 2007. Reasonable Deliberation, Constructive Power and the Struggle for Recognition. In *Recognition and Power*, edited by B. van den Brink and D. Owen. Cambridge: Cambridge University Press.

Lafont, Cristina. 2003. Procedural Justice: Implications of the Rawls-Habermas Debate for Discourse Ethics. *Philosophy and Social Criticism* 29 (2):163-181.

———. 2007. Religion in the Public Sphere: Remarks on Habermas's Conception of Public Deliberation in Postsecular Societies. *Constellations* 14 (2):239-259.

———. 2009. Religion and the Public Sphere: What are the Deliberative Obligations of Democratic Citizenship? *Philosophy and Social Criticism* 35 (1-2):127-150.

Lakatos, Imre. 1978. Falsification and the Methodology of Scientific Research Programmes. In *The Methodology of Scientific Research Programmes: Philosophical Papers Volume 1*, edited by J. Worrall and G. Currie. Cambridge: Cambridge University Press.

Lange, Margaret Meek. 2009. Exploring the Theme of Reflective Stability: John Rawls' Hegelian Reading of David Hume. *Public Reason* 1 (1):75-90.

Larmore, Charles. 1996. *The Morals of Modernity*. Cambridge: Cambridge University Press.

———. 2003. Public Reason. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

———. 2008. *The Autonomy of Morality*. Cambridge: Cambridge University Press.

Lassman, Peter. 2003. Political Theory as Utopia. *History of the Human Sciences* 16 (1):49-62.

Levine, Andrew. 1974. Rawls' Kantianism. *Social Theory and Practice* 3:47-63.

———. 1977. Beyond Justice: Rousseau Against Rawls. *Journal of Chinese Philosophy* 4:123-142.

———. 1993. *The General Will: Rousseau, Marx, Communism*. Cambridge: Cambridge University Press.

Lin, Hsuan-Hsiang. 2008. Beyond the Debate between Cosmopolitanism and Communitarianism: Toward a Hegelian Synthesis.
http://www.allacademic.com//meta/p_mla_apa_research_citation/2/5/3/6/4/page_s253644/p253644-1.php

Lloyd, S.A. 1994. Relativizing Rawls. *Chicago-Kent Law Review* 69:709-735.

Love, Nancy S. 2003. Rawlsian Harmonies: Overlapping Consensus Symphony Orchestra Theory, *Culture and Society* 20 (6):121-40.

Luhmann, Niklas. 2002. I See Something You Don't See. In *Theories of Distinction: Redescribing the Descriptions of Modernity*. Stanford: Stanford University Press.

Lumsden, Simon. 2008. The Rise of the Non-Metaphysical Hegel. *Philosophy Compass* 3 (1):51-65.

Lyons, David. 1989. Nature and Soundness of the Contract and Coherence Arguments. In *Reading Rawls: Critical Studies on Rawls' 'A Theory of Justice'*, edited by N. Daniels. Stanford: Stanford University Press.

Macedo, Stephen. 1990. The Politics of Justification. *Political Theory* 18:280-304.

MacIntyre, Alasdair. 1989. *Whose Justice? Which Rationality?* Notre Dame: University of Notre Dame Press.

_____. 2007. *After Virtue*. 3rd ed. London: Duckworth.

Mandle, Jon. 1997. Justice, Desert, and Ideal Theory. *Social Theory and Practice* 23 (3):399-425.

Manin, Bernard. 1987. On Legitimacy and Political Deliberation. *Political Theory* 15 (3):338-368.

Mardiros, Anthony M. 1952. A Circular Procedure in Ethics. *The Philosophical Review* 61 (2):223-225.

Mason, Andrew. 2004. Just Constraints. *British Journal of Political Science* 34:251-268.

Mason, H.E. 1976. On the Kantian Interpretation of Rawls' Theory. *Midwest Studies in Philosophy 1: Studies in the History of Philosophy*:47-55.

Matan, Ana. 2004. A Well-Ordered Society as a Democratic Community: Alternative Readings of Rawls' Political Theory. *Politička Misao: Croatian Political Science Review* 41 (5):123-133.

Maus, Ingeborg. 2002. Liberties and Popular Sovereignty: On Habermas's Reconstruction of the System of Rights. In *Discourse and Democracy: Essays on Habermas's Between Facts and Norms*, edited by R. von Schomberg and K. Baynes. Albany: State University of New York Press.

McCarthy, Thomas. 1978. *The Critical Theory of Jürgen Habermas*. Cambridge, MA: MIT Press.

_____. 1982. Rationality and Relativism. In *Habermas: Critical Debates*, edited by J. B. Thompson and D. Held. Cambridge: Polity Press.

_____. 1994. Kantian Constructivism and Reconstructivism: Rawls and Habermas in Dialogue. *Ethics* 105 (1):44-63.

_____. 2004. Political Philosophy and Racial Injustice: From Normative to Critical Theory. In *Pragmatism, Critique, Judgment*, edited by S. Benhabib and N. Fraser. Cambridge, MA: MIT Press.

McKinnon, Catriona. 2002. *Liberalism and the Defence of Political Constructivism*. New York: Palgrave.

_____. 2008. *Issues in Political Theory*. Oxford: Oxford University Press.

McMahon, Christopher. 2002. Why There is No Issue between Habermas and Rawls. *The Journal of Philosophy* 99 (3):111-129.

Meckled-Garcia, Saladin. 2008. On the Very Idea of Cosmopolitan Justice: Constructivism and International Agency. *Journal of Political Philosophy* 16:245-271.

Mendus, Susan. 1999. The Importance of Love in Rawls's Theory of Justice. *British Journal of Political Science* 29 (1):57-75.

Mertens, Thomas. 2002. From 'Perpetual Peace' to 'The Law of Peoples': Kant, Habermas and Rawls on International Relations. *Kantian Review* 6:60-84.

Michelman, Frank. 1995. On Regulating Practices with Theories Drawn from Them: A Case of Justice as Fairness. In *Theory and Practice: Nomos XXXVII*, edited by I. Shapiro and J.W. De Cew. New York: New York University Press.

_____. 1997. How Can the People Ever Make the Laws? In *Deliberative Democracy: Essays on Reason and Politics*, edited by J. Bohman and W. Rehg. Cambridge, MA: MIT Press.

_____. 1999. *Brennan and Democracy*. Princeton: Princeton University Press.

Midgaard, Søren Flinch. 2003. Justice, the Basic Structure, and Stability. *Acta Politica* 38:125-145.

Mill, J.S. 1987 [1839]. Bentham. In *John Stuart Mill and Jeremy Bentham: Utilitarianism and Other Essays*, edited by A. Ryan. London: Penguin Books.

———. 1987 [1840]. Coleridge. In *John Stuart Mill and Jeremy Bentham: Utilitarianism and Other Essays*, edited by A. Ryan. London: Penguin Books.

———. 1997 [1831]. The Spirit of the Age. In *Mill: Texts, Commentaries*, edited by A. Ryan. New York: W. W. Norton & Company.

Miller, David. 1999. *Principles of Social Justice*. Cambridge, MA: Harvard University Press.

———. 2002. Two Ways to Think About Justice. *Politics, Philosophy and Economics* 1 (1):5-28.

———. 2008. Political Philosophy for Earthlings. In *Political Theory: Methods and Approaches*, edited by D. Leopold and M. Stears. Oxford: Oxford University Press.

Mills, Charles W. 2005. "Ideal Theory" as Ideology. *Hypatia* 20 (3):165-184.

Mondal, Parthasarathi. 2008. Justice as Fairness: A Quietist Reading of Rawls. *Politikon* 35 (1):107-127.

Moon, J. Donald. 2003. Rawls and Habermas on Public Reason: Human Rights and Global Justice. *Annual Review of Political Science* 6:257-74.

Moore, G.E. 2000 [1903]. The Refutation of Idealism. In *Philosophical Studies*. London: Routledge.

———. 2000 [1919]. External and Internal Relations. In *Philosophical Studies*. London: Routledge.

———. 1993 [1903]. *Principia Ethica*, edited by T. Baldwin. Revised ed. Cambridge: Cambridge University Press.

———. 1993 [1922]. The Conception of Intrinsic Value. In *Principia Ethica*, edited by T. Baldwin. Cambridge: Cambridge University Press.

Mouffe, Chantal. 2005. *The Return of the Political*. London: Verso.

Mulhall, Stephen. 1994. Perfectionism, Politics and the Social Contract: Rawls and Cavell on Justice. *Journal of Political Philosophy* 2 (3):222-239.

———. 1997. Promising, Consent, and Citizenship: Rawls and Cavell of Morality and Politics. *Political Theory* 25 (2):171-192.

Müller, Jan-Werner. 2002. Rawls in Germany. *European Journal of Political Theory* 1 (2):163-179.

———. 2006. Rawls, Historian: Remarks on Political Liberalism's 'Historicism'. *Revue Internationale de Philosophie* 60 (237):327-339.

Munoz-Dardé, Véronique. 1999. Is the Family to be Abolished Then? *Proceedings of the Aristotelian Society, Supplementary* 99 (1):37-56.

Munro, Daniel. 2007. Norms, Motives and Radical Democracy: Habermas and the Problem of Motivation. *The Journal of Political Philosophy* 15 (4):447-472.

Murphy III, Thomas F. 1994. Discourse Ethics: Moral Theory or Political Ethic? *New German Critique* 62:111-135.

Murphy, Liam B. 1999. Institutions and the Demands of Justice. *Philosophy and Public Affairs* 27 (4):251-291.

———. 2003. *Moral Demands in Nonideal Theory*. Oxford: Oxford University Press.

Murphy, Liam and Thomas Nagel. 2002. *The Myth of Ownership: Taxes and Justice*. Oxford: Oxford University Press.

Nagel, Thomas. 1991. *Equality and Partiality*. New York: Oxford University Press.

———. 2000. Getting Personal. *Times Literary Supplement*. 23 June.

Nasr, Waddah N. 1991. On the Proper Function of the Moral Philosopher: Kant and Rawls on Theory and Practice. *Metaphilosophy* 23 (1/2):172-79.

Neal, Patrick. 1997. *Liberalism and its Discontents*. New York: New York University Press.

Neiman, Susan. 2002. *Evil in Modern Thought: An Alternative History of Philosophy*. Princeton: Princeton University Press.

Nelson, William. 1980. The Very Idea of Pure Procedural Justice. *Ethics* 90 (4):502-511.

Neuhouser, Frederick. 1993. Freedom, Dependence and the General Will. *The Philosophical Review* 102 (3):363-95.

———. 2000. *Foundations of Hegel's Social Theory*. Cambridge, MA: Harvard University Press.

———. 2008. *Rousseau's Theodicy of Self-Love*. Oxford: Oxford University Press.

Nielsen, Kai. 1987. Searching for an Emancipatory Perspective: Wide Reflective Equilibrium and the Hermeneutical Circle. In *Anti-foundationalism and Practical Reasoning*, edited by E. Simpson. Edmonton: Academic Printing and Publishing.

Nielsen, Morton Ebbe Juul, ed. 2006. *Political Questions: 5 Questions on Political Philosophy*. Automatic Press/VIP.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

Nussbaum, Martha. 2000. Four Paradigms of Philosophical Politics. *Monist* 83 (4):465-91.

———. 2006. *Frontiers of Justice: Disability, Nationality, Species Membership*. Cambridge, MA: The Belknap Press of Harvard University Press.

O'Neill, Martin. 2009. Liberty, Equality and Property-Owning Democracy. *Journal of Social Philosophy* 40 (3):379-396.

O'Neill, Onora. 1988. Abstraction, Idealization and Ideology in Ethics. In *Ethical Theories and Contemporary Problems*, edited by J.G.D. Evans. Cambridge: Cambridge University Press.

———. 1990. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge: Cambridge University Press.

———. 1992. Vindicating Reason. In *The Cambridge Companion to Kant*, edited by P. Guyer. Cambridge: Cambridge University Press.

———. 1996. *Towards Justice and Virtue: A Constructive Account of Practical Reasoning*. Cambridge: Cambridge University Press.

———. 2000. *Bounds of Justice*. Cambridge: Cambridge University Press.

———. 2003a. Constructivism in Rawls and Kant. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

———. 2003b. Constructivism vs. Contractualism. *Ratio* 16 (4):319-331.

———. 2006. The Method of A Theory of Justice. In *John Rawls. Eine Theorie Der Gerechtigkeit*, edited by O. Höffe. Berlin: Akademie-Verlag.

O'Neill, Shane. 1997. *Impartiality in Context: Grounding Justice in a Pluralist World*. Albany: State University of New York Press.

O'Sullivan, Noël. 2000a. *Political Theory in Transition*. London: Routledge.

———. 2000b. Introduction. In *Political Theory in Transition*, edited by N. O'Sullivan. London: Routledge.

Okin, Susan Moller. 1989. *Justice, Gender, and the Family*. New York: Basic Books.

Olson, Kevin. 2006. *Reflexive Democracy: Political Equality and the Welfare State*. Cambridge, MA: MIT Press.

———. 2007. Paradoxes of Constitutional Democracy. *American Journal of Political Science* 51 (2):330-343.

Outhwaite, William. 1994. *Habermas: A Critical Introduction*. Cambridge: Polity Press.

Owen, David S. 2002. *Between Reason and History: Habermas and the Idea of Progress*. Albany: State University of New York Press.

Pateman, Carole and C. Mills. 2007. *Contract and Domination*. Cambridge: Polity Press.

Peddle, David. 2000. Hegel's Political Ideal: Civil Society, History And *Sittlichkeit*. *Animus* 5.

———. 2004. The Construction of the Secular in Rawls and Hegel: Religion, Philosophy and Public Reason. *Animus* 9.

Peritz, David. 2003. A Diversity of Diversities: Liberalism's Implicit Social Theories. <http://pages.slc.edu/~dperitz/research.php>

———. 2004. Toward a Deliberative and Democratic Response to Multicultural Politics: Post-Rawlsian Reflections on Benhabib's The Claims of Culture. *Constellations* 11 (2):266-90.

Peter, Fabienne. 2007. Democratic Legitimacy and Proceduralist Social Epistemology. *Politics, Philosophy, and Economics* 6 (3):329-353.

Peters, Bernhard. 1996. On Reconstructive Legal and Political Theory. In *Habermas, Modernity and Law*, edited by M. Deflem. London: Sage.

Petersen, Thomas S. and Jesper Ryberg, ed. 2007. *Normative Ethics: 5 Questions*: Automatic Press/VIP.

Pettit, Philip. 1980. *Judging Justice: An Introduction to Contemporary Political Philosophy*. London: Routledge and Kegan Paul.

_____. 1982. Habermas on Truth and Justice. In *Marx and Marxisms*, edited by G.H.R. Parkinson. Cambridge: Cambridge University Press.

_____. 1993. The Contribution of Analytical Philosophy. In *A Companion to Contemporary Political Philosophy*, edited by R. Goodin and P. Pettit. Oxford: Blackwell.

_____. 1996. After Rawls. *Times Literary Supplement*. 17 May.

_____. 1997. *Republicanism: A Theory of Freedom and Government*. Oxford: Clarendon Press.

_____. 1999. Doing Unto Others. *Times Literary Supplement*. 25 June.

_____. 2005. Rawls's Political Ontology. *Politics, Philosophy and Economics* 4 (2):157-174.

_____. 2006. Rawls's Peoples. In *Rawls's Law of Peoples: A Realistic Utopia?*, edited by R. Martin and D. Reidy. Oxford: Blackwell Publishing.

_____. ed. 1991. *Contemporary Political Theory*. New York: Macmillan.

Phillips, Michael. 1985. Reflections on the Transition from Ideal to Non-Ideal Theory. *Nous* 19 (4):551-570.

Pinkard, Terry. 1990. How Kantian was Hegel? *Review of Metaphysics* 43 (4):831-838.

_____. 1999. Virtues, Morality, and *Sittlichkeit*: From Maxims to Practices. *European Journal of Philosophy* 7 (2):217-238.

_____. 2002. *German Philosophy 1760-1860: The Legacy of Idealism*. Cambridge: Cambridge University Press.

Pippin, Robert B. 1995. Hegelianism as Modernism. *Inquiry* 38 (3):305-27.

_____. 1997a. Hegel, Ethical Reasons, Kantian Rejoinders. In *Idealism as Modernism: Hegelian Variations*. Cambridge: Cambridge University Press.

_____. 1997b. Blumenberg and the Modernity Problem. In *Idealism as Modernism: Hegelian Variations*. Cambridge: Cambridge University Press.

_____. 1997c. Hegel, Modernity, and Habermas. In *Idealism as Modernism: Hegelian Variations*. Cambridge: Cambridge University Press.

_____. 1999. *Modernism as a Philosophical Problem: On the Dissatisfactions of European High Culture*. 2nd ed. Oxford: Blackwell.

_____. 2008. *Hegel's Practical Philosophy*. Cambridge: Cambridge University Press.

Plant, Raymond. 1991. *Modern Political Thought*. Oxford: Blackwell.

Pogge, Thomas W. 1981. The Kantian Interpretation of Justice as Fairness. *Zeitschrift für philosophische Forschung* 35 (1):47-65.

_____. 1989. *Realizing Rawls*. Ithaca: Cornell University Press.

_____. 1990. Review of A Treatise of Social Justice, Vol. I: Theories of Justice by Brian Barry. *The Journal of Philosophy* 87 (7):375-84.

_____. 2000. On the Site of Distributive Justice: Reflections on Cohen and Murphy. *Philosophy and Public Affairs* 29 (2):137-169.

_____. 2002. Is Kant's Rechtslehre a 'Comprehensive Liberalism'? In *Kant's Metaphysics of Morals: Interpretive Essays*, edited by M. Timmons. Oxford: Oxford University Press.

_____. 2007. *John Rawls: His Life and Theory of Justice*. Oxford: Oxford University Press.

_____. 2008. Cohen to the Rescue! *Ratio* 21 (4):454-475.

Popper, Karl. 2002 [1959]. *The Logic of Scientific Discovery*. London: Routledge.

Power, Michael. 1993. Habermas and Transcendental Arguments: A Reappraisal. *Philosophy of the Social Sciences* 23 (1):26-49.

———. 1998. Habermas and the Counterfactual Imagination. In *Habermas on Law and Democracy: Critical Exchanges*, edited by M. Rosenfeld. London: University of California Press.

Quine, W.V.O. 2001 [1951]. Two Dogmas of Empiricism. In *Analytic Philosophy: An Anthology*, edited by A.P. Martinich and D. Sosa. Oxford: Blackwell Publishing.

———. 2001 [1960]. *Word and Object*. Cambridge, MA: MIT Press.

———. 2001 [1969]. Epistemology Naturalized. In *Analytic Philosophy: An Anthology*, edited by A.P. Martinich and D. Sosa. Oxford: Blackwell Publishing.

Rajchman, John and C. West ed. 1985. *Post-Analytic Philosophy*. New York: Columbia University Press.

Rasmussen, David M. 2004. Defending reasonability: The centrality of reasonability in the later Rawls *Philosophy and Social Criticism* 30 (5-6):525-540.

Rawls, John. 1950. A Study in the Grounds of Ethical Knowledge: Considered with Reference to Judgments on the Moral Worth of Character. PhD Thesis, Princeton University.

———. 1951. A Review of Stephen Toulmin's *An Examination of the Place of Reason in Ethics*. *Philosophical Review* 60:177-97.

———. 1971/1999. *A Theory of Justice*. Revised ed. Cambridge, MA: Harvard University Press.

———. 1991. John Rawls: For the Record *The Harvard Review of Philosophy* (Spring).

———. 1996. *Political Liberalism*. 2nd ed. New York: Columbia University Press.

———. 1999a. *The Law of Peoples; with, The Idea of Public Reason Revisited*. Cambridge, MA: Harvard University Press.

———. 1999b. *Collected Papers*. Edited by S. Freeman. Cambridge, MA: Harvard University Press.

———. 2000. *Lectures on the History of Moral Philosophy*. Edited by B. Herman. Cambridge, MA: Harvard University Press.

———. 2001a. *Justice as Fairness: A Restatement*. Cambridge, MA: The Belknap Press of Harvard University Press.

———. 2001b. Afterword: A Reminiscence. In *Future Pasts - The Analytic Tradition in Twentieth Century Philosophy*, edited by J. Floyd and S. Shieh. New York: Oxford University Press.

———. 2007. *Lectures on the History of Political Philosophy*. Edited by S. Freeman. Cambridge, MA: The Belknap Press of Harvard University Press.

———. 2009. *A Brief Inquiry into the Meaning of Sin and Faith*. Edited by T. Nagel. Cambridge, MA: Harvard University Press.

Rawls, John and P. van Parijs 2003. Three Letters on the Law of Peoples and the European Union. *Revue de Philosophie Économique* 7 (1):7-20.

Redding, Paul. 2007. *Analytic Philosophy and the Return of Hegelian Thought*. Cambridge: Cambridge University Press.

———. 2008. Georg Wilhelm Friedrich Hegel. In *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*, edited by E. N. Zalta. <http://plato.stanford.edu/entries/hegel/>.

Rehg, William. 1996. Translator's Introduction. In *Between Facts and Norms*. Cambridge, MA: MIT Press.

———. 1998. Against Subordination: Morality, Discourse, and Decision in the Legal Theory of Jürgen Habermas. In *Habermas on Law and Democracy: Critical Exchanges*, edited by M. Rosenfeld. Berkeley and Los Angeles: University of California Press.

———. 2003. Grasping the Force of the Better Argument: McMahon versus Discourse Ethics. *Inquiry* 46 (1):113-133.

Rehg, William and James Bohman. 2001. *Pluralism and the Pragmatic Turn: The Transformation of Critical Theory: Essays in Honour of Thomas McCarthy*. Cambridge, MA: MIT Press.

Reidy, David A. 2008. Introduction. In *John Rawls*, edited by D.A. Reidy. Aldershot: Ashgate.

———. 2009. Rawls's Religion and Justice. http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1352607.

Ricoeur, Paul. 1995. *Oneself as Another*. Chicago: University of Chicago Press.

———. 2000. Is a Purely Procedural Theory of Justice Possible. In *The Just*. Chicago: University of Chicago Press.

Ridge, Michael. 2001. Saving Scanlon: Contractualism and Agent-Relativity. *Journal of Political Philosophy* 9 (4):472-481.

Ripstein, Arthur. 1994. Universal and General Wills: Hegel and Rousseau. *Political Theory* 22 (3):444-467.

Rizvi, Ali Muhammad. 2007. Habermas' Conception of 'Transcendence from Within': An Interpretation. PhD thesis, La Trobe University.

Roberts, Peri. 2007. *Political Constructivism*. London: Routledge.

Robeyns, Ingrid. 2008. Ideal Theory in Theory and Practice. *Social Theory and Practice* 34 (3):341-362.

Rockmore, Tom. 1987. Enlightenment and Reason. *Journal of Philosophy* 84 (11):699-701.

Ron, Amit. 2006. Rawls as a Critical Theorist: Reflective Equilibrium after the 'Deliberative Turn'. *Philosophy and Social Criticism* 32 (2):173-191.

———. 2008. Visions of Democracy in 'Property-Owning Democracy': Skelton to Rawls and Beyond. *History of Political Thought* 29 (1):89-108.

Ronzoni, Miriam. 2009. The Global Order: A Case of Background Injustice? A Practice-Dependent Account. *Philosophy and Public Affairs* 37 (3):229-256.

Rorty, Richard. 1982. Philosophy in America Today. In *Consequences of Pragmatism*. Minnesota: University of Minnesota Press.

———. 1991a. The Priority of Democracy to Philosophy. In *Objectivity, Relativism and Truth*. Cambridge: Cambridge University Press.

———. 1991b. Postmodernist Bourgeois Liberalism. In *Objectivity, Relativism and Truth*. Cambridge: Cambridge University Press.

Rosen, Michael. 1999. Utopia in Frankfurt. *Times Literary Supplement*. 8 October.

Rosenfeld, Michel. 1998a. Can Rights, Democracy, and Justice Be Reconciled through Discourse Theory? Reflections on Habermas's Proceduralist Paradigm of Law In *Habermas on Law and Democracy: Critical Exchanges*, edited by M. Rosenfeld. Berkeley and Los Angeles: University of California Press.

———. 1998b. A Pluralist Critique of Contractarian Proceduralism. *Ratio Juris* 11 (4):291-319.

Rousseau, Jean-Jacques. 1979 [1762]. *Emile: or On Education*. Edited by A. Bloom. New York: Basic Books.

———. 1997 [1755]. Discourse on Political Economy. In *The Social Contract and other later political writings*, edited by V. Gourevitch. Cambridge: Cambridge University Press.

———. 1997 [1762]. The Social Contract. In *The Social Contract and other later political writings*, edited by V. Gourevitch. Cambridge: Cambridge University Press.

Runciman, David. 2006. *The Politics of Good Intentions*. Princeton: Princeton University Press.

Rytina, Steve. 1990. Review of Theories of Justice: A Treatise on Social Justice, Vol. 1 by Brian Barry. *Contemporary Sociology* 19 (6):907-9.

Sandel, Michael J. 1996. *Liberalism and the Limits of Justice*. Cambridge: Cambridge University Press.

Sangiovanni, Andrea. 2007. Global Justice, Reciprocity, and the State. *Philosophy and Public Affairs* 35 (13-39).

———. 2008a. Normative Political Theory: A Flight from Reality? In *Political Thought and International Relations*, edited by D. Bell. Oxford: Oxford University Press.

———. 2008b. Justice and the Priority of Politics to Morality. *Journal of Political Philosophy* 16 (2):137-164.

Saward, Michael. 2000. Less than Meets the Eye: Democratic Legitimacy and Deliberative Theory. In *Democratic Innovation*, edited by M. Saward. London: Routledge.

———. 2002. Rawls and Deliberative Democracy. In *Democracy as Public Deliberation*, edited by M.P. d'Entreves. Manchester: Manchester University Press.

Sayers, S. 1990. Marxism and the Dialectical Method: A Critique of G.A. Cohen. In *Socialism, Feminism and Philosophy: A Radical Philosophy Reader*, edited by S. Sayers and P. Osborne. London: Routledge.

Scanlon, Thomas. 1992. The Aims and Authority of Moral Theory. *Oxford Journal of Legal Studies* 12 (1):1-23.

———. 1998. *What We Owe to Each Other*. Cambridge, MA: Belknap Press of Harvard University Press.

———. 2003a. Rawls on Justification. In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

———. 2003b. Replies. *Ratio* 16 (4):424-39.

Schaar, John. 1974. Reflections on Rawls' Theory of Justice. *Social Theory and Practice* 3 (1):75-100.

Schapiro, Tamar. 2006. Kantian Rigorism and Mitigating Circumstances. *Ethics* 117:32-57.

Scheffler, Samuel. 1994. *The Rejection of Consequentialism*. Revised ed. Oxford: Clarendon Press.

———. 2001. *Boundaries and Allegiances*. Oxford: Oxford University Press.

———. 2003a. What is Egalitarianism? *Philosophy and Public Affairs* 31 (1):5-39.

———. 2003b. Rawls and Utilitarianism In *The Cambridge Companion to Rawls*, edited by S. Freeman. Cambridge: Cambridge University Press.

———. 2003c. Equality as the Virtue of Sovereigns: A Reply to Ronald Dworkin. *Philosophy and Public Affairs* 31 (2):199-206.

———. 2005. The Division of Moral Labour: Egalitarian Liberalism as Moral Pluralism. *Proceedings of the Aristotelian Society, Supplementary* 79:229-253.

———. 2006. Is the Basic Structure Basic? In *The Egalitarian Conscience: Essays in Honour of G.A. Cohen*, edited by C. Sypnowich. Oxford: Oxford University Press.

Scherer, Matthew. 2006. Saint John: The Miracle of Secular Reason. In *Political Theologies: Public Religions in a Post-Secular World*, edited by H. de Vries and L.E. Sullivan. New York: Fordham University Press.

Scheuerman, William E. 1999. Between Radicalism and Resignation: Democratic Theory in Habermas's *Between Facts and Norms*. In *Habermas: A Critical Reader*, edited by P. Dews. Oxford: Blackwell.

Schneewind, J.B. 1963. First Principles and Common Sense Morality in Sidgwick's Ethics. *Archiv fur Geschichte der Philosophie* 45:137-56.

———. 1974. Sidgwick and the Cambridge Moralists. *Monist* 58:371-404.

Schuldenfrei, Richard. 1972. Quine in Perspective. *The Journal of Philosophy* 69:5-16.

Schwartzberg, Melissa. 2008. Voting the General Will: Rousseau on Decision Rules. *Political Theory* 36 (3):403-23.

Schwartzman, Lisa H. 2006. Abstraction, Idealization, and Oppression. *Metaphilosophy* 37 (5):565-588.

Schwarzenbach, Sibyl. 1990. Valuing Ideal Theory: Reflections on Virginia Held's Critique of Rawls. *Metaphilosophy* 21 (1&2):162-178.

_____. 1991. Rawls, Hegel, and Communitarianism. *Political Theory* 19 (4):539-71.

_____. 1994. A Rejoinder to Peter Benson. *Political Theory* 22 (3):501-7.

Scialabba, George 2001. If You're So Egalitarian, How Come You're a Philosopher? *Dissent* (Spring).

Screenivasan, Gopal. 2007. Health and Justice in our Non-ideal World. *Politics, Philosophy and Economics* 6 (2):218-236.

Searle, John R. 1969. *Speech Acts*. Cambridge: Cambridge University Press.

Sen, Amartya. 1980. Equality of What? In *The Tanner Lectures on Human Values I*, edited by S. McMurrin. Salt Lake City: University of Utah Press.

_____. 2006. What do we Want from a Theory of Justice? *The Journal of Philosophy* 103 (5):215-238.

_____. 2009. *The Idea of Justice*. London: Allen Lane.

Sensat, Julius. 2003. Classical German Philosophy and Cohen's Critique of Rawls. *European Journal of Philosophy* 11 (3):314-353.

_____. 2007. Rawlsian Justice and Estrangement: Insights from Hegel and Marx. *Twenty-First Century Papers: On-Line Working Papers from the Center for 21st Century Studies, University of Wisconsin-Milwaukee*. <http://www.uwm.edu/Dept/21st/workingpapers/Sensat.pdf>

Sher, George. 1997. *Approximate Justice: Studies in Non-ideal Theory*. Maryland: Rowman and Littlefield.

Sibley, W.M. 1953. The Rational Versus the Reasonable. *Philosophical Review* 62 (4):554-560.

Sidgwick, Henry. 1981 [1907]. *The Methods of Ethics*. 7th ed. Indianapolis: Hackett.

Silber, John R. 1974. Procedural Formalism in Kant's Ethics. *Review of Metaphysics* 28 (2):197-236.

Singer, Peter. 1974. Sidgwick and Reflective Equilibrium. *Monist* 58:490-517.

Skinner, Quentin. 1998. *Liberty before Liberalism*. Cambridge: Cambridge University Press.

Skinner, Quentin, et al. 2002. Political Philosophy: The View from Cambridge *The Journal of Political Philosophy* 10 (1):1-19.

Smith, Anthony A. 1984. Two Theories of Historical Materialism: G.A. Cohen and Jürgen Habermas. *Theory and Society* 13 (4):513-540.

Smith, Paul. 1998. Incentives and Justice: G.A. Cohen's Egalitarian Critique of Rawls. *Social Theory and Practice* 24 (2).

Smith, S.B. 2007. The Philosopher of Our Times. *New York Sun*. 11 May.

Southwood, Nicholas. 2008. A Deliberative Model of Contractualism. *Politics, Philosophy and Economics* 7 (2):183-208.

Specter, Matthew. 2009. Habermas's Political Thought, 1984-1996: A Historical Interpretation. *Modern Intellectual History* 6 (1):91-119.

Stears, Marc. 2007. Liberalism and the Politics of Compulsion. *British Journal of Political Science* 37:533-553.

Stemplowska, Zofia. 2008. What's Ideal About Ideal Theory? *Social Theory and Practice* 34 (3):319-340.

Stern, Robert. 2006. Hegel's *Doppelsatz*: A Neutral Reading. *Journal of the History of Philosophy* 44 (2):235-66.

_____. 2007. Freedom, Self-Legislation and Morality in Kant and Hegel: Constructivist vs Realist Accounts. In *German Idealism: Contemporary Perspectives*, edited by E. Hammer. London: Routledge.

Sutch, Peter. 2001. *Ethics, Justice and International Relations*. London: Routledge.

Swift, Adam. 1999. Public Opinion and Political Philosophy: The Relation Between Social-Scientific and Philosophical Analyses of Distributive Justice. *Ethical Theory and Moral Practice* 2:337-363.

_____. 2003a. *How Not to be a Hypocrite: School Choice for the Morally Perplexed Parent*. London: Routledge.

———. 2003b. Social Justice: Why Does It Matter What the People Think? In *Forms of Justice: Critical Perspectives on David Miller's Political* edited by D.A. Bell and A. de-Shalit. Oxford: Rowman and Littlefield.

———. 2008. The Value of Philosophy in Non-Ideal Circumstances. *Social Theory and Practice* 34 (3):363-387.

Tampio, Nicholas. 2007. Rawls and the Kantian Ethos. *Polity* 39 (1):79-102.

Tan, Kok-chor. 2004. Justice and Personal Pursuits. *Journal of Philosophy* 7:331-362.

Taylor, Charles. 1975. *Hegel*. Cambridge: Cambridge University Press.

———. 1989. Cross-Purposes: The Liberal-Communitarian Debate. In *Liberalism and the Moral Life*, edited by N.L. Rosenblum. Cambridge, MA: Harvard University Press.

———. 1996. The Motivation behind a Procedural Ethics. In *Kant and Political Philosophy: The Contemporary Legacy*, edited by R. Beiner and W.J. Booth. New Haven: Yale University Press.

Tribe, Laurence H. 1985. *Constitutional Choices*. Cambridge, MA: Harvard University Press.

Tucker, R.C. 1969. *The Marxian Revolutionary Idea*. New York: W.W. Norton.

van Mill, David. 1996. The Possibility of Rational Outcomes from Democratic Procedures. *Journal of Politics* 58 (3):734-752.

Vergés-Gifra, Joan. 2006. Why Rawls Should Have Resisted the Hermeneutic Temptation. *Journal of Social Philosophy* 37 (4):584-603.

Vincent, Andrew. 2004. *The Nature of Political Theory*. Oxford: Oxford University Press.

Voorhoeve, Alex. 2001. Kant on the Cheap: Thomas Scanlon Interviewed. *The Philosophers' Magazine* 16:29-30.

———. 2005. Incentives and Principles for Individuals in Rawls's Theory of Justice. *Ethics and Economics* 3.

Waldron, Jeremy. 1993. Theoretical Foundations of Liberalism. In *Liberal Rights*. New York: Cambridge University Press.

———. 1999. *Law and Disagreement*. Oxford: Oxford University Press.

Walzer, Michael. 1981. Philosophy and Democracy. *Political Theory* 9 (3):379-399.

———. 1984. *Spheres of Justice*. New York: Basic Books.

Watkins, Erik and W. Fitzpatrick. 2002. O'Neill and Korsgaard on the Construction of Normativity. *The Journal of Value Inquiry* 36:349-367.

Weithman, Paul. 1994. Liberalism and the Political Character of Political Philosophy. In *The Liberalism-Communitarianism Debate*, edited by C.F. Delaney. Lanham: Rowman and Littlefield.

———. 2007. Review of Samuel Freeman, Justice and the Social Contract: Essays on Rawlsian Political Philosophy. *Notre Dame Philosophical Reviews*. <http://ndpr.nd.edu/review.cfm?id=10405>.

———. 2009. John Rawls and the Task of Political Philosophy. *The Review of Politics* 71:113-125.

Wellmer, Albrecht. 1991. Ethics and Dialogue. In *The Persistence of Modernity*. Cambridge: Polity Press.

Wenar, Leif. 2004. The Unity of Rawls's Work. *The Journal of Moral Philosophy* 1 (3):265-75.

———. 2008. John Rawls. In *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, edited by E. N. Zalta. <http://plato.stanford.edu/archives/fall2008/entries/rawls/>.

White, Morton. 2002. *A Philosophy of Culture: The Scope of Holistic Pragmatism*. Princeton: Princeton University Press.

White, Stephen K. 1987. *The Recent Work of Jürgen Habermas: Reason, Justice and Modernity*. Cambridge: Cambridge University Press.

Williams, Andrew. 1998. Incentives, Inequality, and Publicity. *Philosophy and Public Affairs* 27 (3):225-247.

Williams, Bernard. 1981. Internal and External Reasons. In *Moral Luck*. Cambridge: Cambridge University Press.

_____. 1993. *Ethics and the Limits of Philosophy*. London: Fontana Press.

_____. 1997. Forward to Basics. In *Equality*, edited by J. Franklin. London: IPPR.

_____. 2005a. Realism and Moralism in Political Theory. In *In the Beginning Was the Deed*, edited by G. Hawthorn. Princeton: Princeton University Press.

_____. 2005b. In the Beginning Was the Deed. In *In the Beginning Was the Deed*, edited by G. Hawthorn. Princeton: Princeton University Press.

_____. 2005c. Modernity and the Substance of Ethical Life. In *In the Beginning Was the Deed*, edited by G. Hawthorn. Princeton: Princeton University Press.

_____. 2006. Political Philosophy and the Analytic Tradition. In *Philosophy as a Humanistic Discipline*, edited by A.W. Moore. Princeton: Princeton University Press.

Williams, Howard. 1999. Kant, Rawls, Habermas and the Metaphysics of Justice. *Kantian Review* 3:1-17.

Williams, Robert R., ed. 2001. *Beyond Liberalism and Communitarianism*. Albany: State University of New York Press.

Wingenbach, Ed. 1999. Unjust Context: The Priority of Stability in Rawls's Contextualized Theory of Justice. *American Journal of Political Science* 43 (1):213-232.

Wolff, Jonathan. 1998. Fairness, Respect and the Egalitarian Ethos. *Philosophy and Public Affairs* 27 (2):97-122.

_____. 2007. Equality: The Recent History of an Idea. *Journal of Moral Philosophy* 4 (1):125-136.

Wolff, Jonathan and A. de-Shalit. 2007. *Disadvantage*. Oxford: Oxford University Press.

Wolff, Robert Paul. 1977. *Understanding Rawls*. Princeton: Princeton University Press.

Wolin, Sheldon S. 1996. The Liberal/Democratic Divide: On Rawls's Political Liberalism. *Political Theory* 24 (1):97-119.

Wood, Allen W. 1999. Autonomy as the Ground of Morality. *O'Neil Memorial Lectures*. <http://www.stanford.edu/~allenw/webpapers/Autonomy.doc>.

Yates, Melissa. 2007. Rawls and Habermas on Religion in the Public Sphere. *Philosophy and Social Criticism* 33 (7):880-891.

Žižek, Slavoj. 2008. *In Defense of Lost Causes*. London: Verso.