

Republics of Commitments

Pluralism from the Individual to the Liberal State

A DISSERTATION SUBMITTED TO
CARNEGIE MELLON UNIVERSITY FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY.

AUGUST 20, 2010

David Emmanuel Gray

*Department of Philosophy
Carnegie Mellon University*



Τὸ ζήτημα ὃ ἐπιχειροῦμεν οὐ φαῦλον ἀλλ' ὀξύ βλέποντος, ὡς ἐμοὶ φαίνεται. ἐπειδὴ οὖν ἡμεῖς οὐ δεινοί, δοκεῖ μοι, ἦν δ' ἐγώ, τοιαύτην ποιήσασθαι ζήτησιν αὐτοῦ, οἷανπερ ἂν εἰ προσέταξέ τις γράμματα σμικρὰ πόρρωθεν ἀναγνῶναι μὴ πάνυ ὀξύ βλέπουσιν, ἔπειτά τις ἐνενόησεν, ὅτι τὰ αὐτὰ γράμματα ἔστι που καὶ ἄλλοθι μείζω τε καὶ ἐν μείζονι, ἔρμαιον ἂν ἐφάνη, οἶμαι, ἐκεῖνα πρῶτον ἀναγνόντας οὕτως ἐπισκοπεῖν τὰ ἐλάττω, εἰ τὰ αὐτὰ ὄντα τυγχάνει.

The investigation we're undertaking is not an easy one but requires keen eyesight. Therefore, since we aren't clever people, we should adopt the method of investigation that we'd use if, lacking keen eyesight, we were told to read small letters from a distance and then noticed that the same letters existed elsewhere in a larger size and on a larger surface. We'd consider it a godsend, I think, to be allowed to read the larger ones first and then to examine the smaller ones, to see whether they really are the same.

Plato, *Republic* II.368c–d

Contents

Introduction	xi
1 Pluralism and the Liberal State	1
1.1 The Fundamental Aim of the Liberal State	2
1.2 The Fact of Reasonable Pluralism	6
1.2.1 Systematicity and Comprehensive Doctrines	6
1.2.2 The Inevitability of Reasonable Pluralism	13
1.3 Liberal Approaches to Political Legitimacy	18
1.4 Responsive Impartiality and State Agency	28
2 Liberal Democratic Proceduralism	35
2.1 Procedural Justifications	38
2.2 Epistemic Conceptions of Democracy	44
2.2.1 Bentham's Utilitarian Justification of Democracy	45
2.2.2 Rousseau and the Condorcet Jury Theorem	49
2.3 Contractualist Defenses of Democracy	55
2.3.1 Constitutional Contractualism	58
2.3.2 Deeply Democratic Contractualism	67
2.4 Purely Procedural Liberal Democracy	73
3 Social Choice Theory and Liberal Democracy	81
3.1 The Default Rationality Thesis	83
3.2 Social Choice Theory and Arrow's Theorem	88
3.3 Bayesian Rationality	92

3.4	The Rationality Postulates	95
3.5	The Condorcet Rule and Domain Restriction	101
3.5.1	Multidimensional Structuration	107
3.6	Responsive Impartiality for Pooling Functions	110
3.6.1	Cardinal Valuations and Interpersonal Comparisons	116
3.7	Implications for Liberal Democracy	124
4	Reconciling Asymmetries in Reasoning	129
4.1	The Liberal State as Rational Agent	132
4.2	The Symmetry Between the Individual and the Liberal State	137
4.3	Impossibility Theorems and the Symmetry Thesis	143
5	Agreement and Legitimacy	151
5.1	Maximality	153
5.2	Admissibility	167
5.2.1	Some Difficulties With Admissibility	178
5.3	Agreement on Reasons Versus Agreement on Options	188
5.4	Agreement in the Liberal State	195
5.5	Concluding Thoughts	200
	Bibliography	201

List of Figures

3.1	The three ways that two options may be ordered.	84
3.2	Two value functions with identical orderings of the options. . . .	85
3.3	A standard decision matrix.	93
3.4	A profile of individual value orderings that has a majority ranking an option over the plurality winner.	96
3.5	A profile illustrating Condorcet's paradox.	98
3.6	A profile leading the Condorcet and narrow Borda rules to pro- duce identical value orderings.	99
3.7	A profile leading the Condorcet and narrow Borda rules to pro- duce conflicting value orderings.	100
3.8	Illustrating how changing the available options influences the results of the narrow Borda rule.	101
3.9	Example of a single-peaked profile.	103
3.10	Example of a multi-peaked profile.	104
3.11	Disaggregating a multi-peaked profile into two single-peaked profiles.	109
3.12	An example of how the broad Borda rule violates independence of irrelevant alternatives.	113
3.13	Two alternative profiles of value functions for Tipper and Albert. . .	117
3.14	An additional profile for Tipper and Albert.	121
4.1	Dr. Bailey's evaluation of the efficacy of three treatment alloca- tions on three patients.	140

4.2	The three different ways in which Dr. Webber's values and commitments rank three different treatments for a single patient. . .	141
4.3	Comparing the default rationality thesis with the extended rationality thesis.	148
4.4	Three coalitions and their respective evaluations of three different health care policies that the liberal state may adopt.	148
5.1	Barry and Michelle's profile of value orderings for selecting a power plant.	156
5.2	George and Laura's profile of value orderings for selecting a power plant.	157
5.3	William and Hillary's profile of value orderings for selecting a power plant.	161
5.4	Two profiles of value orderings concerning the selection of a power plant, where the first is lexically prior to the second. . . .	163
5.5	Two alternative profiles of value functions for William and Hillary. .	165
5.6	Two extensions of William and Hillary profile of value orderings taking into account a neutral probability mixture.	167
5.7	The graph of all the possible weighted averages of William and Hillary's value functions when they agree that one of the options is second worst.	170
5.8	The graph of all the possible weighted averages of William and Hillary's value functions when they agree that one of the options is second best.	171
5.9	A decision matrix for selecting a power plant.	174
5.10	The graph of all the possible weighted averages of two Bayesian value functions when they disagree over probabilities.	175
5.11	William and Hillary's rival decision matrices for constructing a power plant.	177
5.12	Two profiles of value functions concerning the selection of a power plant, where the first is lexically prior to the second.. . .	178
5.13	Dick and Jane's evaluations of the consequences for treating a patient.	181
5.14	Dick and Jane's decision matrix for treating a patient.	182
5.15	Meredith and Christina's decision matrix for treating a patient. .	184
5.16	Three different probability distributions for four states of affairs. .	184
5.17	Schematic of Mario and Luigi's growing fields.	185
5.18	Five different crop allocations for Mario and Luigi's fields. . . .	186
5.19	Illustration of the doctrinal paradox.	191
5.20	Iterated illustration of the doctrinal paradox from figure 5.19. . .	192

5.2.1	Derek and Owen’s Bayesian value functions assessing two treatments for a patient.	192
5.2.2	Derek and Owen’s rival decision matrices for the effectiveness of two treatments with two states of affairs.	193
5.2.3	Derek and Owen’s rival decision matrices for the effectiveness of two treatments with three states of affairs.	194

Introduction

PROCEDURAL approaches to political legitimacy have become increasingly popular amongst liberals. According to such an approach, the legitimacy of a state decision is primarily derived from the processes followed in order to make that decision and not from the quality of the decision itself. The processes that liberals have in mind are typically those found within a system of democratic institutions. These electoral and legislative procedures are supposed to allow the state's constitutive members to reach legitimately binding agreements on how the state should exercise its power.

In response to this trend in liberalism, this dissertation has two goals. The first is to argue that formal results in the social decision sciences seriously question the viability of these procedural approaches. The second goal, however, is to argue that these results do not necessitate complete abandonment of procedural approaches but require liberals to more carefully specify the sorts of agreements that legitimacy requires and what procedures may reliably produce such agreements. I ultimately agree that procedures are important for liberalism but they require more careful understanding and defense.

In order to carry this project out, this dissertation is divided into five chapters. In the first two chapters, I outline a fairly standard account of liberalism while arguing that it contains a serious tension that liberal democratic institutions are supposed to resolve. The second chapter in particular seeks to explain the various ways in which the procedural aspects of democracy are particularly appealing in this respect. In the third chapter, I then present Arrow's famous impossibility result along with other formal result from social choice theory suggesting that in certain situations there simply is no correct procedure for legitimate outcomes that the liberal state may employ. The fourth chapter responds to this concern by

maintaining that there exists a symmetry between how an individual person and the liberal state make decisions, and that insights into individual decision making sheds light on how it may be done by a group of people in a liberal democracy. Finally, in the fifth chapter, I move in this direction by presenting an account of the sort of agreements between individuals that legitimacy ought to be grounded upon. Ultimately, procedures like voting remain an important aspect of liberal democracy for authorizing state action, though not necessarily for the reasons that many liberal democrats endorse.

CHAPTER ONE

Pluralism and the Liberal State

LIBERALISM is typically thought to require that legitimate exercises of power by the state be agreed upon and authorized by those subject to them. Societal pluralism, however, suggests the difficulty for agreement on how the state should respond to controversial issues. The concern is whether there exist any principles for justifying state action that are appropriately neutral or impartial between the various values and beliefs of the state's constitutive members.¹ As a result, it remains unclear what constitutes a legitimate decision when there is no unanimous consent. This concern has led contemporary liberals to increasingly advocate a system of democratic electoral and legislative processes for making decisions, as opposed to a set of process-independent principles for assessing decisions, that the state should employ.

In this chapter, I present a brief outline of my understanding of liberalism and a tension that its commitments to legitimacy and toleration generate. In particular, I begin by identifying what I believe to be the fundamental aim of the liberal state (section 1.1). I then argue that pursuing this aim has two important consequences for the liberal state. First, it invites and even encourages a particularly disconcerting form of societal pluralism within the liberal state (section 1.2). Second, it posits a fairly demanding condition for the legitimate exercise of state power (section 1.3).

¹Throughout, I refer to the state's 'constitutive members' or its 'people' instead of its 'citizens'. Doing the latter presumes an account of citizenship and providing such an account lies beyond the scope of this dissertation. I therefore speak of the state's constitutive members, since doing so has a fairly intuitive meaning. Otherwise, I defer to the discussion of citizenship in Robert A. Dahl, *Democracy and Its Critics* (New Haven, CT: Yale University Press, 1989), pp. 119–131. Dahl maintains that the state's constitutive members consists of all mentally fit, adult residents governed by the state. Such a definition is sufficient for my purposes.

Together, these two issues suggest that legitimacy requires some form of state neutrality—or what I prefer to call responsive impartiality—which comes into tension with the need for the state to also pursue a consistent and coherent political agenda (section 1.4). In the next chapter, I then focus on how democratic processes are supposed to ease this tension.

1.1 The Fundamental Aim of the Liberal State

The core values of the liberal tradition are usually expressed in terms of liberty and equality.² Liberalism insists that all persons hold in common a fundamental interest to develop and exercise their autonomy to the fullest of its potential. By autonomy, I refer to a person's capacity to freely form, revise, and pursue her own conception of the good life and of worthwhile ends on the basis of what she regards as good reasons.³ According to liberalism then, a person should not only

²While the label of 'liberalism' certainly applies to a whole constellation of disparate positions, I nevertheless believe my outline of liberalism captures the central ideas of most conventional liberal views. My presentation of liberalism is especially influenced here by Ronald Dworkin, 'Liberalism', in *Public and Private Morality*, ed. by Stuart Hampshire (Cambridge: Cambridge University Press, 1978): 113–143, 'To Each His Own', *New York Review of Books* 30, 6 (14th April 1983): 4–6, 'What is Equality? – Part 3: The Place of Liberty', *Iowa Law Review* 73, 1 (October 1987): 1–54, 'Foundations of Liberal Equality', in *The Tanner Lectures on Human Values*, ed. by Grethe B. Peterson, vol. 11 (Salt Lake City: University of Utah Press, 1990): 1–119; John Rawls, 'Kantian Constructivism in Moral Theory', *Journal of Philosophy* 77, 9 (September 1980): 515–572, 'The Basic Liberties and Their Priority', in *The Tanner Lectures on Human Values*, ed. by Sterling M. McMurrin, vol. 3 (Salt Lake City: University of Utah Press, 1982): 1–87, 'Social Utility and the Primary Goods', in *Utilitarianism and Beyond*, ed. by Amartya K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982): 159–186; Bruce Ackerman, *Social Justice in the Liberal State* (New Haven, CT: Yale University Press, 1980); Jeremy Waldron, 'Theoretical Foundations of Liberalism', *Philosophical Quarterly* 37, 147 (April 1987): 127–150; Thomas Nagel, *Equality and Partiality* (Oxford: Oxford University Press, 1991); Christine M. Korsgaard, 'Commentary on G. A. Cohen and Amartya Sen', in *The Quality of Life*, ed. by Martha C. Nussbaum and Amartya K. Sen (Oxford: Clarendon Press, 1993): 54–61; Brian Barry, *A Treatise on Social Justice*, vol. 2: *Justice as Impartiality* (Oxford: Oxford University Press, 1995); Martha C. Nussbaum, *Sex and Social Justice* (Oxford: Oxford University Press, 1999); and Elizabeth S. Anderson, 'What Is the Point of Equality?', *Ethics* 109, 2 (January 1999): 287–337.

³John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 19, refers to this as one of the two 'moral powers' possessed by all persons, the 'capacity for a conception of the good'. According to *ibid.*, p. 74, all persons have a 'higher-order' interest in developing and exercising this capacity. See also Rawls, 'Kantian Constructivism', p. 525 and 'Social Utility and the Primary Goods', pp. 164–165, where this is characterized as a 'highest-order' interest. I do not mention the other moral power Rawls discusses, the 'capacity for a sense of justice', which is the capacity to understand, apply, and act from fair terms of cooperation with others, though it remains implicit in my discussion. Also see the discussions on the value of autonomy in Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), pp. 369–429 and Thomas Hurka,

be able to rationally pursue whatever she values but also be free to choose and even reassess what she should value in light of new reasons and circumstances she may encounter. While this emphasis on respecting individual autonomy may assume many forms, it remains an important focal point in the works of classical liberals such as John Locke, Immanuel Kant, and John Stuart Mill as well as those of contemporary liberals such as Robert Nozick and John Rawls. These accounts suggest that all people should have significant liberty in discovering, evaluating, and pursuing whatever it is that gives their lives value, meaning, and purpose.⁴

Liberalism recognizes and accepts that most people may affirm a conception of the good life that is commonly accepted by their peers, either on due reflection or simply as a matter of custom; but it also demands toleration for those who experiment with more idiosyncratic ways of life. Liberalism is suspicious of threats, physical violence, and other tactics that force a person to accept a particular conception of how she ought to live or what she should value. According to this view, a person should be free to take personal responsibility for such important matters. Liberalism maintains that coercion is normally wrong because it attempts to force a person to conform to something without appeal to what counts as good reason—by *that person's* lights—justifying why she should conform to it. This is what is meant by the assertion that coercion is disrespectful to autonomy. Liberalism therefore requires that people respect the autonomy of others by restraining themselves, except under special circumstances, from interfering with the beliefs and actions of others.⁵

From its commitment to individual liberty, liberalism derives a corollary regarding the equal moral worth of all persons, regardless of their station in society. This corollary asserts that all competent adults equally have the power to develop and exercise their autonomy and that this entitles each to equal respect and consideration.⁶ This is not to presume that everyone possesses equal talent, nor necessarily

⁴‘Why Value Autonomy?’, *Social Theory and Practice* 13, 3 (Fall 1987): 361–382.

⁵As noted in Barry, *Justice as Impartiality*, p. 129, liberalism may be interpreted as regarding the development and exercise of autonomy as components of a conception of the good life, but, as Barry continues, it is a ‘second-order conception’ insofar as it does not specify the full contents of the good life. Persons are left to come to their own conclusions on this matter, provided that they each freely make up their own minds.

⁶It is for this reason that liberalism is equally suspicious of Marxist and socialist theories as it is of ones demanding allegiance to religion or tradition. For more on this, see Stephen Holmes, *The Anatomy of Antiliberalism* (Cambridge, MA: Harvard University Press, 1993).

⁷As Anderson, ‘What Is the Point of Equality?’, pp. 314–315, argues, freedom and equality are mutually reinforcing ideals because equals are not normally subject to coercion, oppression, or other limitations on their freedom by others. This is in contrast with those who insist that liberalism’s understanding of freedom and equality, such as Friedrich A. Hayek, *The Constitution of Liberty* (Chicago: University of Chicago Press, 1960), pp. 85–102 and Milton Friedman and Rose D. Friedman, *Free to Choose: A Personal Statement* (Orlando, FL: Harcourt, 1980), pp. 128–149. Following

to demand that they all enjoy the same socioeconomic status. What it means is that differences in characteristics such as talent, social class, gender, race, or sexual orientation cannot justify hierarchical distinctions of moral worth. In other words, no one's autonomy is more important than anyone else's; no one is inherently subservient to another. As a result, liberalism seeks to abolish most forms of oppression and intolerance, such as exploiting the labor of a group for the exclusive benefit of another or marginalizing certain people from useful participation in and contribution to society.⁷

Given that liberalism regards all persons as free and equal, this raises the question of how liberalism justifies the formation of political society, which involves binding individuals to a recognized set of governing institutions that establish and enforce rules of association. This is especially relevant when addressing the state, as it claims exclusive control over a particular geographic region and asserts ultimate authority within that region. While liberalism does not affirm the naturalness or inherent value of subjecting a person to such an authority, it nevertheless maintains that there are prudential and moral reasons for relinquishing a certain amount of one's liberty in order to enter into the liberal state.

As do most political theories, liberalism justifies the state's existence on the basis of the security it provides its people. Life outside of all political society is thought to be uncertain and fragile because individual survival depends crucially upon interacting with one's neighbors. The state, on the other hand, can coordinate its constitutive members in order to create a reliable and effective military, police, and civil justice system for securing liberty by providing protection against arbitrary coercion and oppression. Not only is this arrangement desirable for those subject to it, but it also provides the conditions necessary for these people to live free and equal, allowing them all to develop and exercise individual autonomy. For persons are hardly free and equal if some are allowed to abuse others without reason, or if some are exploited and marginalized by others. The liberal state is supposed to secure its members against threats like these as well, and so helps

Waldron, 'Theoretical Foundations of Liberalism', p. 129n6, the slogan is perhaps best understood as 'equal liberty for all', a statement consistent with the views of libertarians such as Hayek and Friedman. Even so, non-libertarian liberals may draw implications from this slogan that go far beyond what these libertarians are willing to endorse. For instance, preventing oppression and intolerance may require the elimination of large disparities in wealth through economic redistribution—a process flatly rejected by libertarians. Doing so, however, is only justified according to liberalism if the process can be shown to be a necessary means for the equal respect and promotion of individual autonomy.

⁷Exploitation and marginalization—along with powerlessness, cultural imperialism, and violence—are among the 'five faces of oppression' identified by Iris Marion Young, *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990), pp. 39–65.

fulfill liberalism's commitment to liberty and equality.⁸

This approach to justifying the liberal state reveals how liberalism conceives of the state's proper organization. The fundamental aim of the liberal state, as I understand it, is to secure for its constitutive members social and political arrangements for equally respecting and promoting their individual autonomy. On this account, the state must adopt certain liberal institutions such as the rule of law, the entitlement of persons to basic control over their own bodies and property, and the liberties of conscience and thought.⁹ These are typically extended to include a variety of freedoms associated with religion, expression, movement, private property, political participation, the press, and assembly that apply equally to all the state's constitutive members and cannot be abridged according to a person's talent, social class, gender, race, or sexual orientation. In the liberal state, these institutions have a certain, though not necessarily absolute, priority over other public and private goods that the state and its members may pursue. In addition, the state may provide certain social resources—such as a system for public education and some arrangement for market transactions—that each of its members may equally draw upon in order to more effectively develop and pursue a conception of the good life.

Pursuing the fundamental aim of the liberal state has two important consequences. First, when the state protects individual liberties, its people have the freedom to determine for themselves what ends they value and the lives they intend to pursue. It seems inevitable that these ends and lives will eventually diverge and create societal pluralism. As Mill points out, liberal institutions allow 'experiments in living', whereby 'persons of genius' may discover whole new ends and ways of life.¹⁰ Second, achieving this aim undoubtedly requires the state to exercise its powers of coercion over its constitutive members. This raises the concern that the liberal state paradoxically, to paraphrase Rousseau, forces its people to be free.¹¹ In the next two sections, I explain each of these issues in turn. Following that, the tension raised for liberalism will become apparent.

⁸Of course, many liberals extend the potential benefits of the state to include provisions such as a market system and recognized currency, public infrastructure supporting roads and utilities, hospitals, schools, and so on. Coordinating these outside all forms of political society would be extraordinarily difficult and unlikely to succeed.

⁹In the next chapter, I add to this list of institutions the electoral and legislative processes normally associated with a liberal democracy.

¹⁰*On Liberty* (1859), chapter III in volume 18 of John Stuart Mill, *Collected Works of John Stuart Mill*, ed. by J. M. Robson (Toronto: University of Toronto Press, 1963–1991), pp. 260–275.

¹¹*On the Social Contract* (1762), book I, chapter vii in Jean-Jacques Rousseau, *On the Social Contract, with Geneva Manuscript and Political Economy*, ed. by Roger D. Masters, trans. by Judith R. Masters (New York: St. Martin's Press, 1978), p. 55.

1.2 The Fact of Reasonable Pluralism

Contemporary liberals usually express a concern with the impact of societal pluralism on the liberal state. Unfortunately, though, there is not yet a clear account of the different forms of pluralism, their differing influences, and the different ways in which the liberal state may respond to them. A useful starting point on this issue, though, is the later work of John Rawls, where he insists that pluralism is unavoidable for the modern liberal state.¹² He calls this the ‘fact of reasonable pluralism’.¹³ The pluralism Rawls has in mind is not simply the result of cultural diversity or a wide variety of interest groups and other organizations with competing agendas. It is a reasonable pluralism of comprehensive doctrines. Rawls, along with many other liberals, express concerns with such a pluralism, and the importance of ensuring that the liberal state remains neutral between the competing comprehensive doctrines endorsed by its constitutive members. While I return to the issue of state neutrality later (in section 1.4), I first outline what Rawls means by a ‘comprehensive doctrine’ while arguing this account must be extended to recognize that there are a variety of features that make doctrines comprehensive in different respects. This sets the stage for my position later (in section 5.4) that the liberal state may endorse a doctrine that is comprehensive in certain of these respects, though not comprehensive in other respects. In this section, I also present why Rawls believes that a ‘reasonable’ pluralism over such doctrines is unavoidable in the liberal state (section 1.2.2). I detail a significant portion of Rawls’ particular understanding of pluralism because it is well-formulated in comparison with other liberals, and I believe that the impact of such pluralism for contemporary liberalism cannot be overstated.

1.2.1 Systematicity and Comprehensive Doctrines

Rawls envisions that an unavoidable consequence of liberal institutions involves the state’s members affirming a variety of ‘comprehensive doctrines’. As Rawls explains it:

[A doctrine] is comprehensive when it includes conceptions of what is of value in human life, and ideals of personal character, as well as

¹²John Rawls, ‘Justice as Fairness: Political not Metaphysical’, *Philosophy and Public Affairs* 14, 3 (Summer 1985): 223–251, ‘The Idea of an Overlapping Consensus’, *Oxford Journal of Legal Studies* 7, 1 (Spring 1987): 1–25, ‘The Priority of Right and Ideas of the Good’, *Philosophy and Public Affairs* 17, 4 (Autumn 1988): 251–276, ‘The Domain of the Political and Overlapping Consensus’, *New York University Law Review* 64, 2 (May 1989): 233–255, *Political Liberalism*, and ‘The Idea of Public Reason Revisited’, *University of Chicago Law Review* 64, 3 (Summer 1997): 765–807.

¹³Rawls, *Political Liberalism*, p. 36.

ideals of friendship and of familial and associational relationships, and much else that is to inform our conduct, and in the limit to our life as a whole. A conception is fully comprehensive if it covers all recognized values and virtues within one rather precisely articulated system.¹⁴

As such, a comprehensive doctrine usually takes a stand on a wide variety of issues, such as economic prosperity, property, social status, health, education, sexual behavior, abortion, religion, security, equality, liberty, social justice, art and culture, the environment, the treatment of non-human animals, and whatever else a person might care about. Some of these positions may be commonly accepted while others may not. The doctrine itself may be religious, moral, or philosophical in nature. What is essential about a comprehensive doctrine, however, is that it attempts to organize its views on a range of issues into a consistent and coherent whole.

Beyond the diversity of their particular content, Rawls is not entirely clear on distinguishing the various ways in which doctrines can be comprehensive. Furthermore, there has been little discussion by others concerning differences that may be quite significant for the proper understanding of state neutrality. I maintain that doctrines can be more or less comprehensive according to a variety of dimensions, three of which should be made more explicit: inclusiveness, generality, and systematicity. Inclusiveness concerns the comprehensiveness of content, and a doctrine is inclusive to the extent that it encompasses a broad range of goods and values. Generality concerns the scope of evaluation and the more general a doctrine, the wider the range of subjects to which it applies. Systematicity concerns the degree to which a doctrine incorporates its associated goods and values into a single, unified system weighing and ordering them with relative precision.¹⁵

To illustrate, classical utilitarianism is usually presented as the comprehensive doctrine par excellence.¹⁶ First, the principle of utility is inclusive because it subsumes a wide variety of goods and values, including conceptions of personal

¹⁴ibid., p. 13.

¹⁵Unlike the other dimensions, systematicity is largely left implicit by Rawls, though it does emerge from time to time. For instance, it is seen at ibid., p. 145, when Rawls describes a 'partially comprehensive' doctrine for his 'model case' as 'not systematically unified' and 'pluralistic' because it leaves 'all values to be balanced against one another, either in groups or singly, in particular kinds of cases'. This dimension also appears at ibid., pp. 154–156, when Rawls addresses the objection that a fully comprehensive doctrine is necessary to 'order the many conflicts of justice that arise in public life'.

¹⁶I have in mind Bentham's *An Introduction to the Principles of Morals and Legislation* (1789); Mill's *Utilitarianism* (1861); and Henry Sigwick, *The Methods of Ethics*, 7th ed. (London: Macmillan, 1907).

virtue and social justice.¹⁷ This principle also has a very general scope, applying to individual conduct, social and political institutions, and international relations.¹⁸ Furthermore, the principle of utility is very systematic, as it delineates a single measure, calling it 'utility', for comparing and weighing various goods and values while resolving any conflicts that might arise between them.¹⁹ In its classical formulations, utility typically refers to a single principal good such as pleasure or happiness, which in turn prioritizes and orders all other goods and values. In recent formulations, utility takes on a more abstract, all-things-considered, measure such as 'preference satisfaction' for ordering all goods and values.²⁰ No matter its particular form, the striking feature of utilitarianism is its systematic monism: the principle of utility commensurates any apparent value pluralism, such as competing goals and moral commitments or the holdings of goods across different individuals, into a single assessment.²¹ It is precisely this sort of systematicity that motivates Rawls' argument that 'utilitarianism does not take seriously the distinction between persons'.²²

Despite this protestation against utilitarianism, some may still maintain that systematicity itself may be an attractive, and perhaps even necessary, feature for

¹⁷Especially chapters IV–V of Mill's *Utilitarianism* (1861) in volume 10 of Mill, *Collected Works*, pp. 234–259.

¹⁸Rawls, *Political Liberalism*, p. 260, notes that the distinction between act and rule utilitarianism does not necessarily undermine this claim. Regardless of whether its evaluative focal point is acts or rules, the principle of utility is, directly or indirectly, assessing an extremely wide scope of social forms and individual activities.

¹⁹For instance, in *Utilitarianism* (1861), chapter II in volume 10 of Mill, *Collected Works*, p. 226, Mill declares that 'if utility is the ultimate source of moral obligations, utility may be invoked to decide between them when their demands are incompatible'.

²⁰This is apparent in the utilitarianism of John C. Harsanyi, 'Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking', *Journal of Political Economy* 61, 5 (October 1953): 434–435, 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility', *Journal of Political Economy* 63, 4 (August 1955): 309–321, and 'Bayesian Decision Theory and Utilitarian Ethics', *American Economic Review* 68, 2 (May 1978): 223–228.

²¹Ambiguities abound in the philosophical literature concerning what it means for two or more things to be 'commensurable'. See Ruth Chang, 'Introduction', in *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997): 1–34, for an attempt at some clarification on this and related concerns. As I use it, two or more things are commensurable with respect to a given standard if and only if they are measurable by that standard. To commensurate two or more things is to specify such a measure. I do not, however, assume that this measure must necessarily involve a *numerical* scale; rather, the measure need only allow for an *ordinal*—or rank ordering—scale sufficient for making comparisons. To illustrate, a measure comparing the performance of any two runners in a given race need only denote the faster and the slower person (or whether they tied). It need not be more precise, say, by specifying the number of seconds separating their respective running times.

²²John Rawls, *A Theory of Justice* (Cambridge, MA: Belknap Press of Harvard University Press, 1971), p. 27.

a doctrine. That is, I believe that an unsystematic doctrine is typically thought somehow deficient, problematic, and in need of revision. In fact, systematicity may be thought more attractive or necessary for a doctrine than the other two dimensions of comprehensiveness. It does not seem too problematic, for instance, if a doctrine is non-inclusive or non-general: the doctrine simply does not apply with respect to certain values, goods, or subjects. If the adherent of an exclusive or localized doctrine must make a choice between options that solely involve values, goods, or subjects outside of its range of application, this person may then claim indifference without any apparent difficulty. Each of the options seems just as good as any of the others. However, I believe that it is generally thought that a doctrine borders on the inconsistent, incoherent, ad hoc, and arbitrary when it is not systematic with respect to the values, goods, and subjects to which it does apply. This is because an unsystematic doctrine fails to provide a complete account for weighing, ordering, or otherwise comparing its associated values and goods, which in turn raises the concern that such a doctrine may at times make conflicting or incompatible demands leading to ad hoc and arbitrary decision making. If this is correct, a common assumption appears to be that systematicity is a necessary feature for a doctrine to even be rational, or just reasonable, an idea to which I return shortly (in section 1.2.2).

To illustrate this desirability of systematicity, suppose Jones values a life of both patriotism and pacifism, but she does not systematically delineate the relative importance of these two commitments. The worry is that if Jones' nation goes to war and she must choose between enlisting in the army and conscientious objection, her commitments come into opposition, generating a conflict without immediate resolution.²³ In this situation, Jones is hardly indifferent between her options, for whatever she chooses, she fails to satisfy one of her commitments. Yet since her view is unsystematic, neither option is, from this perspective, better than the other. This seems to call for further reflection by Jones in order to determine how to weigh patriotism and pacifism when they conflict. Increasing the systematicity of her view in this or some other way is apparently necessary. If Jones' view remains unsystematic, however, it is difficult to understand how she may make a decision that is consistent and coherent with all her commitments as opposed to one arbitrarily chosen. Reasoning along these lines supports what I take to be a common assumption concerning systematicity: asystematicity in a person's view is a sort of embarrassment that should be avoided as much as possible. Regardless, a commitment to systematicity does not mean that individuals should never modify

²³This example is from John Dewey and James H. Tufts, *Ethics*, revised edition (New York: Henry Holt and Company, 1932), p. 174, though a similar one appears in Jean-Paul Sartre's 'Existentialism is a Humanism', Jean-Paul Sartre, 'Existentialism', in *Existentialism and Human Emotions*, trans. by Bernard Frechtman (New York: Philosophical Library, 1957): 9–51, pp. 24–25.

their views in light of new reasons and experiences—say by introducing new values or goods—but that, on the whole, individuals should strive for rational unity in their views even as it may be so modified.

At this point, I should emphasize that I do not fully endorse the necessity of systematicity for rational decision making. Starting in section 4.3, I explicitly challenge it. Right now, I am only suggesting its appeal. Indeed, the need for systematicity appears within many views, including some that purport to eschew it. For instance, consider Ronald Dworkin's and Robert Nozick's respective theories of individual rights as 'trumps' over maximizing utility or other possible ends.²⁴ At times, these accounts appear to accord rights absolute weight, where competing considerations can never infringe or override any right. Such a theory may be both inclusive and general by incorporating a wide range of values for regulating a variety of social arrangements and interactions and yet not be very systematic. It certainly establishes an ordering between rights and non-rights, immediately resolving any conflict between the two, but it may not specify any hierarchy of rights. Values are therefore ordered to a certain extent, though not completely or with great precision. To take a simple example, if second-hand smoke is indeed harmful, it seems that a smoker's right to smoke in public conflicts with a non-smoker's right not to inhale harmful substances. Without weighing or ordering rights, conflicts between them remain unresolved.²⁵ Of course, this theory is similarly silent on conflicts between non-rights as well. The unwillingness to engender conflicts of these sorts best explains why Dworkin and Nozick each weaken the requirement that rights have absolute weight.²⁶ In other words, resolving or otherwise avoiding these types of conflicts apparently require increasing the theory's systematicity.

Other inclusive and general doctrines may claim to be even less systematic than the theory of rights-as-trumps. Such a doctrine may recognize a wide plurality of values but either not provide a weighing or an ordering of these values, or it may explicitly deny that such comparisons are possible. In short, a doctrine may purport to abstain from resolving conflicts between different values and goods. Michael

²⁴The phrase is from Ronald Dworkin, *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1978), p. xi. Also see *ibid.*, pp. 364–368 and Ronald Dworkin, 'Is There a Right to Pornography?', *Oxford Journal of Legal Studies* 1, 2 (Summer 1981): 177–212, pp. 199–206. A similar theory of rights as constraints is presented by Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 26–53.

²⁵For more on this issue see Joseph Raz, 'Professor Dworkin's Theory of Rights', *Political Studies* 26, 1 (March 1978): 123–137; Thomas Nagel, 'Libertarianism without Foundations', *Yale Law Journal* 85, 1 (November 1975): 136–149; and Judith Jarvis Thomson, 'Some Ruminations on Rights', *Arizona Law Review* 19, 1 (1977): 45–60.

²⁶For instance, see Dworkin, *Taking Rights Seriously*, p. 366 and Nozick, *Anarchy, State, and Utopia*, pp. 30n, 41, *Philosophical Explanations*, (Cambridge, MA: Belknap Press of Harvard University Press, 1981), pp. 479–504.

Walzer's theory of goods is one example of a doctrine like this.²⁷ It is extremely inclusive and general in scope, encompassing a wide array of different goods—health, wealth, power, piety, and knowledge—that together encompass every aspect of a person's life. Nevertheless, this theory strives towards extremely asystematicity because it maintains that these goods fall within distinct and autonomous 'spheres' of evaluation that are incommensurable by moral and political standards.²⁸ For instance, 'money is inappropriate in the sphere of ecclesiastical office', whereas 'piety should make for no advantage in the marketplace'.²⁹ It is true that within some spheres there may be a complete weighing or ordering of goods, but there is no privileged sphere or primary good that sets the definitive ordering over all other spheres and goods. This means that different spheres may provide conflicting assessments of the same goods or the same actions, and this theory denies that it has a definitive resolution.³⁰ Even so, Walzer's position ultimately backs away from this asystematicity when it apparently concedes that within a given tradition, culture, or society, there exist means for arriving at a more or less systematic understanding of the proper weighing and ordering of these spheres.³¹ If my

²⁷Michael Walzer, *Spheres of Justice: A Defense of Pluralism and Equality* (New York: Basic Books, 1983).

²⁸As with 'commensurable' (see n. 21) there are ambiguities in the philosophical literature concerning what it means for two things or more things to be 'incommensurable'. As I will use it, two things are incommensurable with respect to a given standard if and only if that standard cannot provide a measure for comparing them. For instance, the respective works of Mozart and Michelangelo may be incommensurable with respect to overall artistic merit, as it seems that neither is strictly better than the other and their achievements are hardly equal. Given their radically different artistic media, there is simply no single, 'all-things-considered' standard for comparing the overall artistic merit of their respective works. This need not preclude claiming, however, that the music of Antonio Salieri is decidedly worse on this scale when compared to the works of either Mozart or Michelangelo, respectively. In that case, the fact that Salieri and Mozart are commensurable, as are Salieri and Michelangelo, need not imply that Mozart and Michelangelo are commensurable.

²⁹Walzer, *Spheres of Justice*, p. 10.

³⁰To illustrate further: 'if the religious uses of bread were to conflict with its nutritional uses—if the gods demanded that bread be baked and burned rather than eaten—it is by no means clear which use would be primary', *ibid.*, p. 8. Just before this, Walzer does entertain the idea that the nutritional uses of bread may outweigh any religious ones when there is a food shortage, though he then adds that 'even there, we can't be sure'.

³¹For instance, see the discussion of health care at *ibid.*, pp. 86–90, or the discussion of the relativity of justice at *ibid.*, pp. 312–316. Similar rejections of asystematicity occur within a variety of theories advocating value pluralism when they suggest some single, overarching scale for weighing the various values the respective theory recognizes. See, for instance, Charles Taylor, 'The Diversity of Goods', in *Utilitarianism and Beyond*, ed. by Bernard Williams and Amartya K. Sen (Cambridge: Cambridge University Press, 1982): 129–144, 'Leading a Life', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 170–183; Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge: Cambridge University Press, 1985); James Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance*

understanding of Walzer's theory is correct, then it rejects the existence of some universally privileged weighing or ordering of spheres while granting that a person must herself have some such mode of comparison, which I presume this person is supposed to derive from her understanding of her society's particular values and traditions. Such value pluralism may endorse asystematicity in the abstract, but this support evaporates in practice.

Apart from distinguishing systematicity from other notions of comprehensive-ness, I also want to separate issues of systematicity from those involving teleological specificity. By the latter, I refer to the extent to which a doctrine specifies a particular conception of the good life or of worthwhile ends that people should pursue. I mention this because it may be tempting to assume that a fully systematic doctrine must have teleological specificity. This is certainly the case with many religious doctrines that take a comprehensive account of the nature of God and God's relationship with all of creation to derive an understanding of the good life as freely conforming to God's specific plan for humanity. Such specificity, though, need not encompass all systematic doctrines. The theory of rights-as-trumps may be otherwise rendered systematic. Doing so need not, however, entail that a person should pursue any particular conception of the good life, even though such a theory may not be compatible with all such conceptions. Even utilitarianism may lack teleological specificity, as Kant suggests when arguing that 'although every human being wishes to attain [happiness], he can still never say determinately and consistently with himself what he really wishes and wills'.³² That is, claiming that happiness, whether understood as pleasure or preference satisfaction, should be the end of all human action may only provide apparent specificity. Unless a distinction is made between the types of pleasures, happiness, or satisfaction that a person ought to seek, utilitarianism leaves it to each individual to discover for herself what exactly it is that makes her happy.³³

(Oxford: Clarendon Press, 1986), James Griffin, 'Incommensurability: What's the Problem?', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 35–51; Michael Stocker, *Plural and Conflicting Values* (Oxford: Clarendon Press, 1990); Joseph Raz, 'Incommensurability and Agency', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 110–128, *Engaging Reason: On the Theory of Value and Action*, (Oxford: Oxford University Press, 1999); and Chang, 'Introduction', 'Putting Together Morality and Well-Being', in *Practical Conflicts: New Philosophical Essays*, ed. by Peter Baumann and Monika Betzler (Cambridge: Cambridge University Press, 2004): 118–158.

³²*Groundwork of the Metaphysics of Morals* (1785), section II [4:418] in Immanuel Kant, *Practical Philosophy*, ed. and trans. by Mary J. Gregor (Cambridge: Cambridge University Press, 1996), p. 70. This point is made when Kant argues against happiness as an objective foundation for morality.

³³Of course, Mill's distinction between higher and lower pleasures in chapter II of *Utilitarianism* (1861), volume 10 of Mill, *Collected Works*, pp. 210–213, attempts to do just this.

This also cuts the other way, however, for a doctrine may have teleological specificity while not being systematic. Theories of professional morality often have this structure. Physicians, for instance, are often thought to have a special ‘therapeutic obligation’ to safeguard the health and wellbeing of their patients. Such role-related doctrines are not inclusive or general because they only specify ideals for a particular professional relationship or other domain of life, and even within that domain, they may not be very systematic. Your doctor may, for example, rightfully ask to see you naked in the context of a physical exam, but she cannot expect this of you in a non-medical setting. Furthermore, the therapeutic obligation may involve commitments to both patient autonomy and patient health, for example, without organizing these commitments in a systematic fashion, leaving it to the individual medical practitioner to make any hard choices. Despite lacking comprehensiveness, role-based moralities do specify particular ends that their respective professions should pursue.

Given these different dimensions for understanding comprehensive doctrines, different types of pluralism might appear in society. For instance, the members of an all-Christian state might share a common conception of the good life as living in accord with God’s plan for humanity, but adopt different doctrines concerning the systematic weighing and comparing of particular goods and values. The impact of this pluralism on the liberal state is almost certainly different from the pluralism that emerges when the state’s constitutive members adopt various doctrines that have no commonalities along any of these dimensions. The latter type of pluralism seems more complicated, and yet inevitable for the modern liberal state. The nature of this pluralism should be especially disconcerting for liberalism because, as Rawls persuasively argues, it remains even when confined to ‘reasonable’ doctrines.

1.2.2 The Inevitability of Reasonable Pluralism

A pluralism of comprehensive doctrines need not arise solely from aggressive, closed-minded selfishness and sectarianism; from errors in reasoning; or from a propensity of persons to express their individuality.³⁴ If these were the only

³⁴Hobbes’ *Leviathan* (1651), part I, chapter xiii in volume 3 of Thomas Hobbes, *The English Works of Thomas Hobbes of Malmesbury*, ed. by Sir William Molesworth (London: John Bohn, 1839–1845), pp. 110–116, forcefully argues that the fear of death and the competitive desire for commodious living—both heightened by the desire for glory—naturally set individuals against one another. Nevertheless, in the right situation, aggressively self-interested people should cooperate, as the folk theorem applied to iterated prisoner dilemmas shows, see Ken Binmore, *Fun and Games: A Text on Game Theory* (Lexington, MA: D. C. Heath, 1992), pp. 366–377. For a fairly exhaustive, but accessible, survey of the many biases and errors that regularly plague human reasoning, see Reid Hastie and Robyn M. Dawes, *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making* (Thousand Oaks, CA: Sage, 2001). Stuart Hampshire, *Innocence*

sources of pluralism, a liberal state might effectively curb it through education, community activities, promotion of civic virtue, and other similar policies. Doing this might move especially noxious doctrines like social Darwinism, neo-Nazism, and religious fundamentalism—along with views that are irrational or simply crazy—to the fringes of society where they may have little influence. Whether it is desirable or not for the liberal state to suppress comprehensive doctrines like these, doing so is not enough to eliminate pluralism. Many liberal views now accept, following Rawls, that there would still remain a pluralism of ‘reasonable’ comprehensive doctrines.³⁵

Rawls is ‘deliberately loose’ in specifying what makes a comprehensive doctrine reasonable in hopes of avoiding ‘the danger of being arbitrary and exclusive.’³⁶ He does suggest that a reasonable doctrine is consistent and coherent in organizing its evaluations concerning the goods and values it encompasses, that it attempts to resolve conflicts between these goods and values when they conflict, and that it evolves over time in response to critical reflection, new evidence and information, and whatever other reasons that it judges to be relevant.³⁷ Given this description, it certainly appears that Rawls holds the position that systematicity is necessary for a doctrine to be reasonable. For without it, a doctrine seems unable to organize its commitments in the way Rawls prescribes for a reasonable doctrine. Consequently, an unreasonable doctrine runs the risk of leading its adherent to making ad hoc and arbitrary decisions.

Beyond these, Rawls notes that reasonable doctrines ‘are the doctrines that reasonable citizens affirm’ and he extensively develops the idea of a reasonable person.³⁸ Among other things, Rawls claims that reasonable people (1) have comparable intellectual powers of reasoning, thought, judgment, drawing inferences, and weighing evidence;³⁹ (2) are committed to reciprocity in the sense that they propose terms of cooperation they sincerely believe are justified by reasons that other reasonable people may accept;⁴⁰ (3) recognize that other reasonable people

and Experience (Cambridge, MA: Harvard University Press, 1989), pp. 32–38, 41–48, posits that human imagination and a drive towards individuality leads people to adopt different doctrines.

³⁵Rawls, *Political Liberalism*, pp. 47–88, 235–239, which in turn draws on Thomas Nagel, ‘Moral Conflict and Political Legitimacy’, *Philosophy and Public Affairs* 16, 3 (Summer 1987): 215–240 and Joshua Cohen, ‘Moral Pluralism and Political Consensus’, in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 270–291.

³⁶Rawls, *Political Liberalism*, p. 59.

³⁷*ibid.*, p. 59.

³⁸*ibid.*, p. 36.

³⁹*ibid.*, p. 55.

⁴⁰*ibid.*, pp. 48–54. Also see John Rawls, ‘Introduction to the Paperback Edition’, in *Political Liberalism* (New York: Columbia University Press, 1996), pp. xlv–xlvii.

may endorse differing comprehensive doctrines;⁴¹ (4) agree to adhere to basic liberal institutions, provided that other reasonable people may be relied on to do so as well;⁴² and (5) may disagree over the particular arrangements and applications of those institutions.⁴³

While feature (1) is fairly self-explanatory as a mild idealization of reasonableness, Rawls' defense of the other claims is somewhat terse, so it is worth briefly fleshing out.⁴⁴ In discussing the reasonableness of persons, Rawls says that 'the reasonable is public' and that 'insofar as we are reasonable, we are ready to work out the framework for the public social world'.⁴⁵ This suggests that people behave unreasonably when they desire to literally force their views on others; whereas they are reasonable when they attempt to reason with each other, that is, they are prepared to offer, and even accept, terms that they actually expect other reasonable people can accept. Even those who view political society simply as a *modus vivendi* held together by a fragile balance of power between various conflicting factions must appreciate, unless they believe they have overwhelming force backing them up, that the best way to improve their circumstances is to propose terms to others in this fashion. The reasonableness of the commitment to reciprocity might then be justified on these grounds.

Feature (3) contains an important admission: disagreements can exist even among reasonable people. That is, people with similar levels of intellectual ability who are also genuinely cooperative and open-minded may still endorse competing reasonable comprehensive doctrines. In defending this, Rawls points to several 'burdens of judgment' that inevitably lead reasonable people to disagree with one another.⁴⁶ Four such burdens are representative:

[1] The evidence—empirical and scientific—bearing on the case is conflicting and complex, and thus hard to assess and evaluate.

[2] Even where we agree fully about the kinds of considerations that are relevant, we may disagree about their weight, and so arrive at different judgments. . . .

[3] Often there are different kinds of normative considerations of

⁴¹This is a consequence the 'fifth general fact' of political sociology in *Political Liberalism*, p. 58.

⁴²*ibid.*, p. xviii.

⁴³*ibid.*, pp. 229–230.

⁴⁴For instance, in defending the commitment to reciprocity, *ibid.*, pp. 48, 51, asserts that 'everyday speech' recognizes this distinction and that unreasonable people verge on the 'psychopathic'. Without further explanation, Rawls' understanding of reasonableness risks either begging the question or becoming a sterile technical definition.

⁴⁵*ibid.*, p. 53.

⁴⁶*ibid.*, pp. 54–58. They are called 'burdens of reason' in Rawls, 'Domain of the Political and Overlapping Consensus', pp. 235–238.

different force on both sides of an issue and it is difficult to make an overall assessment.

[4] Finally, . . . any system of social institutions is limited in the values it can admit so that some selection must be made from the full range of moral and political values that might be realized. . . . In being forced to select among cherished values, or when we hold to several and must restrict each in view of the requirements of the others, we face great difficulties in setting priorities and making adjustments. Many hard decisions may seem to have no clear answer.⁴⁷

Burdens like these prevent reasonable people from converging onto a single reasonable comprehensive doctrine. This is not to deny that the burdens of reason may be overcome eventually; but for the foreseeable future, they appear unavoidable. As a result, reasonable pluralism, according to Rawls, 'is rooted in the difficulties of exercising our reason under the normal conditions of human life'.⁴⁸

Feature (4) is perhaps the most contentious of these features and it does seem to be a mere stipulation by Rawls. Nevertheless, feature (4) retains some plausibility in light of features (2) and (3). For consider a society dominated by people who affirm various comprehensive doctrines compatible with feature (2). In light of feature (3), they recognize that they can reasonably disagree over these doctrines and that further debate, even when organized in terms on which reasonable people can agree, will not necessarily lead to agreement on a single doctrine. Preferring to live in peace or worrying that they might lose any civil war, it seems plausible that these people might judge adopting basic liberal institutions to be the best option available to them.⁴⁹ Feature (4) therefore entails that agreeing to adhere to liberal

⁴⁷Rawls, *Political Liberalism*, pp. 56–57. These are burdens a, b, e, and f, respectively.

⁴⁸'Domain of the Political and Overlapping Consensus', p. 239.

⁴⁹*Political Liberalism*, pp. 146–149, suggests this, which is a combination of two rather distinct arguments. Following Brian Barry, 'How Not to Defend Liberal Institutions', in *Liberalism and the Good*, ed. by R. Bruce Douglass, Gerald R. Mara and Henry S. Richardson (London: Routledge, 1990): 44–58, these two arguments are the argument from social peace (liberal institutions are the best way for a divided society to avoid civil war) and the argument from prudence (liberal institutions are the best way for each party to avoid the worst of all possibilities, being forced to adhere to a set of alien beliefs). As *ibid.*, p. 47, notes, these involve two assumptions: first, the parties 'attribute a high value to social peace' over civil war and, second, that there is 'actually . . . a balance of forces within the society'. Barry's assertions to the contrary aside, the first premise hardly seems controversial. This premise plays a role in the largely illiberal political philosophies of Plato and Hobbes. For instance, in Plato's *Republic*, book V [462a–b] in Plato, *Complete Works*, ed. by John M. Cooper and D. S. Hutchinson (Indianapolis, IN: Hackett, 1997), p. 1089, civil war is the greatest evil 'we can mention' for a society, and in Plato's *Laws*, book I [628a–d] in *ibid.*, pp. 1322–1323, it is suggested that most people do place a high value on civil peace. Meanwhile, Hobbes' *Leviathan* (1651), part I, chapters xiii–xv in Hobbes, *English Works*, pp. 110–147, argues that this preference explains why people leave the state of nature and form states. Regarding the second premise, this

institutions does not require adopting a particular comprehensive doctrine. With that implication, feature (5) seems inevitable: if different doctrines support liberal institutions, the burdens of judgment should lead to reasonable disagreements over how to arrange these institutions and apply them to particular cases.

This last point is worth emphasizing, for it is a somewhat disconcerting aspect of reasonable pluralism vis-à-vis liberalism. Liberal institutions may be valued by a variety of different comprehensive doctrines for a variety of different reasons. First-order liberal doctrines, for instance, endorse liberal institutions along the lines that I presented in section 1.1. According to these doctrines, the commitment to liberty and equality demands protecting basic liberties.⁵⁰ Second-order liberal doctrines, on the other hand, do not share this commitment but still support liberal institutions as means to satisfy their other commitments.⁵¹ Such a doctrine might be moderately self-interested and therefore portray these institutions simply as a necessary part of a *modus vivendi*. On this account of reasonableness, a variety of first- and second-order liberal doctrines may all be perfectly reasonable and yet have conflicting positions on precise content of liberal institutions, their application, and how to weigh and order them should they ever conflict.⁵² Fundamental aspects of liberalism itself are therefore open to a certain amount of reasonable disagreement.

Rawls posits that this pluralism of reasonable comprehensive doctrines is an inevitable feature of the liberal state, which he calls the fact of reasonable pluralism.⁵³ He adds that only the sustained and oppressive exercise of power by the state may suppress pluralism, and doing this seems decidedly antithetical to liberalism.⁵⁴ As I noted at the end of section 1.1, however, the liberal state cannot refrain from all exercises of power if it is to secure for its constitutive members arrangements that equally respect and promote their autonomy. This leads to

is certainly a social contingency, but one that now appears to be firmly entrenched in the public culture of most societies.

⁵⁰First-order liberal doctrines include most traditional liberal theories, such as those presented in Locke's *Second Treatise of Government* (1690); Kant's *The Metaphysics of Morals* (1797); Rawls, *Theory of Justice*; and Nozick, *Anarchy, State, and Utopia*.

⁵¹A classic example of a second-order liberal doctrine appears in Bentham's *Principles of the Civil Code* (1843), where he argues that personal and political liberties are means to security, which in turn provides greater happiness. See part I, chapters ii, vii–vii in volume 1 of Jeremy Bentham, *The Works of Jeremy Bentham*, ed. by John Bowring (London: Simpkin, Marshall, and Company, 1838–1843), pp. 302, 307–309.

⁵²To take a simple example, Cass Sunstein, *Democracy and the Problem of Free Speech* (New York: The Free Press, 1993) and Joshua Cohen, 'Freedom of Expression', *Philosophy and Public Affairs* 22, 3 (Summer 1993): 207–263, apparently disagree over whether freedom of expression only protects political speech or extends to literary, artistic, scientific, and other forms of expression.

⁵³Rawls, *Political Liberalism*, p. 36.

⁵⁴*ibid.*, p. 37, calls this the 'fact of oppression'.

the issue of specifying the conditions in which the liberal state may legitimately exercise its coercive power.

1.3 Liberal Approaches to Political Legitimacy

There are two stark realities of politics: first, the state must exercise its power in order to achieve its aims, whatever they may be, and, second, every use of this power is coercive. In the case of the liberal state, coercion is necessary in order to achieve its fundamental aim, which is to provide sociopolitical arrangements for equally respecting and promoting the autonomy of its members. This gives rise to what might initially appear to be a contradiction: liberalism is committed to the free exercise and development of individual autonomy while also endorsing coercion of individuals to conform to certain liberal institutions. At the very least, the liberal state promises to frustrate the oppressive aspirations of neo-Nazis, religious fundamentalists, and other intolerant groups. Additionally, the fact of reasonable pluralism suggests that it may at times also frustrate the aspirations of some of its more tolerant members.

The concern is then how people—including the intolerant—meaningfully retain their individual liberty when a powerful authority effectively binds them to its laws and decisions. With the modern state a dominant feature of contemporary life, most people are born subjugated to this sort of authority and they have little hope of ever completely escaping it. Given that liberalism does not regard this as a natural situation, it must account for how a person can be beholden to an authority that she may never have freely chosen to obey.

To address this concern, liberalism must explain the circumstances that effectively bind the state's constitutive members to its laws and decisions, especially when these laws and decisions may be at odds with a particular individual's conception of the good life. In other words, liberalism must provide the grounds for the liberal state's 'right to rule'.⁵⁵ There are two ways of understanding what this 'right' involves. On the one hand, it may mean that each of the state's constitutive members has an obligation or duty to comply with its laws and decisions; on the other hand, it may mean that the state may permissibly coerce compliance with them. Following a common convention, I refer to the former as political

⁵⁵See David Copp, 'The Idea of a Legitimate State', *Philosophy and Public Affairs* 28, 1 (Winter 1999): 3–45 and Christopher W. Morris, 'Natural Rights and Political Legitimacy', *Social Philosophy and Policy* 22, 1 (January 2005): 314–329. Though helpful, this phrase may be misleading: *X* having a right to *Y* is often thought to logically correlate with an obligation on certain specified agents to provide *X* with *Y*, for instance, or to not interfere with *X* obtaining and holding *Y*. If this is assumed, then it seems that differentiating political obligation and political legitimacy becomes a distinction without a difference.

obligation (or authority) and the latter as political legitimacy.⁵⁶

It is currently rather common in political philosophy to insist that political legitimacy need not correlate with political obligation.⁵⁷ This position asserts that the state may rightly exercise its power to enforce compliance with certain laws and decisions even though its constitutive members may not have an obligation to obey them.⁵⁸ This is an important issue, but I do not address it because most contemporary liberals appear primarily concerned with explaining how the exercise of state power is consistent with a commitment to liberty and equality.⁵⁹ Therefore, I

⁵⁶'Legitimacy' is often used as a modifier applied to states, meaning that a legitimate state has acquired a certain authority over its constitutive members in virtue of the way in which the state came to have power over them. If a military coup, for instance, overthrows a democratically elected government and establishes a dictatorship, the resulting state is usually judged illegitimate. As will become apparent shortly, my account of political legitimacy primarily involves exercises of power by the state (meaning that the state may rightfully coerce its constitutive members in the specified circumstances), rather than the state itself (meaning that the state has acquired a more general sort of authority—such as to the loyalty and allegiance of its people—owing to how it came to power). This is a distinction recognized by Locke, for example, in the *Second Treatise of Government* (1690) where he argues that the state acquires the power to rule only when its constitutive members unanimously agree to give it this power. Once this is done, however, Locke grants that majority rule is sufficient to make a particular law or decision enforceable by the state. See sections 95–99 in volume 4 of John Locke, *The Works of John Locke, in Nine Volumes*, 12th ed. (London: C. and J. Rivington, et al., 1824), pp. 394–396.

⁵⁷See Robert Ladenson, 'In Defense of a Hobbesian Conception of Law', *Philosophy and Public Affairs* 9, 2 (Winter 1980): 134–59; Rolf Sartorius, 'Political Authority and Political Obligation', *Virginia Law Review* 67, 1 (February 1981): 3–17; Waldron, 'Theoretical Foundations of Liberalism'; Kent Greenawalt, *Conflicts of Law and Morality* (Oxford: Oxford University Press, 1987), pp. 47–61; William A. Edmundson, 'Legitimate Authority without Political Obligation', *Law and Philosophy* 17, 1 (January 1998): 43–60; Cynthia A. Stark, 'Hypothetical Consent and Justification', *Journal of Philosophy* 97, 6 (June 2000): 313–334; Allen Buchanan, *Justice, Legitimacy, and Self-Determination* (Oxford: Oxford University Press, 2004), pp. 233–241; and David M. Estlund, *Democratic Authority* (Princeton: Princeton University Press, 2008), pp. 40–64, 117–135, among many others.

⁵⁸The classic example from M. B. E. Smith, 'Is There a Prima Facie Obligation to Obey the Law?', *Yale Law Journal* 82, 5 (April 1973): 950–976, involves whether to stop at a stop sign at a deserted highway intersection. In this case, it seems difficult to understand why a person has an obligation to obey the stop sign since no plausible harm could come from failing to come to a full stop. Even so, it seems much less difficult to understand why a police officer may issue a ticket to a person caught ignoring that stop sign.

⁵⁹In light of the broad demands of the modern state on the loyalties and allegiances of its constitutive members, I am sympathetic to the position defended by A. John Simmons, *Moral Principles and Political Obligations* (Princeton: Princeton University Press, 1979), 'Justification and Legitimacy', *Ethics* 109, 4 (July 1999): 739–771, 'Political Obligation and Authority', in *The Blackwell Guide to Social and Political Philosophy*, ed. by Robert L. Simon (Oxford: Blackwell, 2002): 17–37, and 'The Duty to Obey and Our Natural Moral Duties', in *Is There a Duty to Obey the Law? For and Against*, ed. by Christopher Heath Wellman and A. John Simmons (Cambridge: Cambridge University Press, 2005): 91–196. Simmons maintains that political obligation and political legitimacy cannot be easily untangled. His reasoning begins with the observation that

focus on political legitimacy and leave it as an open question concerning the extent to which this involves political obligation. For the remainder of this section, I lay out a brief taxonomy of liberal approaches to political legitimacy in order to isolate what I believe to be the dominant approach in contemporary liberalism.

Generally speaking, liberalism adopts at least two approaches to political legitimacy. The authoritative approach suggests that the state's exercise of power in a given situation is legitimate only when done to equally preserve, protect, and promote the autonomy of all its constitutive members, while the consensual approach insists that the state's exercise of power in a given situation is legitimate only when the state's constitutive members collectively authorize it. I emphasize immediately that this is not supposed to be too severe a dichotomy, for a liberal theory might employ aspects of both authoritative and consensual approaches.⁶⁰

The authoritative approach contends that the commitment to liberty and equality requires the state to adopt basic liberal institutions and that coercing compliance with them is fundamentally consistent with an equal concern for individual autonomy. According to this approach, liberty is not license to do whatever one pleases, but about respecting everyone's—not just one's own—capacity to freely make fundamental decisions about the value, meaning, and purpose of one's life.⁶¹ In this manner, the commitment to equality specifies the proper scope of the commitment to individual liberty. Therefore, while liberal institutions provide a person with an expansive amount of freedom, this freedom is constrained, usually with some form of the 'harm principle', which prohibits one from causing harm to others against their will.⁶² The authoritative approach thus stresses a person's entitlement to choose freely what ends to pursue, but it also means that there are limits to how people may treat each other when pursuing their respectively chosen ends.⁶³ The state may then, according to this view, legitimately exercise its power

the modern state claims to have sole responsibility for enforcing compliance with its institutions, policies, laws, and decisions. This prohibits other states and vigilante groups from exercising this authority, thus implying that political legitimacy involves placing an obligation of non-competition on all other agents concerning enforcement within that state's territory. Simmons concludes that establishing the obligation of non-competition would most likely involve the same sort of arguments that would also establish a moral general obligation of obedience.

⁶⁰Examples might include the liberal theories of Locke and Kant. Both thinkers seem to employ an authoritative approach to justify the need for the state to adopt basic liberal institutions. They may then be understood as taking a consensual approach to explain how the state should exercise its power in a manner that is consistent with these institutions.

⁶¹For instance, see Ronald Dworkin, 'Did Mill Go Too Far?', *New York Review of Books* 21, 17 (31st October 1974), and 'What is Equality? – Part 3'.

⁶²The classic statements of this principle are found in Locke's *Second Treatise of Government* (1690), section 6 in volume 4 of Locke, *Works*, p. 341 and Mill's *On Liberty* (1859), chapter I in volume 18 of Mill, *Collected Works*, p. 223.

⁶³This is often expressed by the principle of the priority of the right over the good, as found in

to enforce these limits.

Certainly this enforcement may require the state to interfere with the actions of neo-Nazis and religious fundamentalists, but the authoritative approach maintains that doing so still respects and promotes individual autonomy. The argument is that the autonomy of members of intolerant groups like these is respected insofar as they remain free to hold their views and even free to try and convince others to willingly accept them. It is when the intolerant adopt coercive or oppressive tactics infringing on the liberty of others that the state must intervene. Thus, coercion by the state is permissible when it interferes with those who attempt to interfere with others. This echoes Kant's argument that while hindering someone's freedom is wrong, it is nonetheless permissible to hinder the freedom of someone who is herself hindering the freedom of another. In this way, as Kant explains, the '*hindering of a hindrance to freedom . . . is consistent with freedom*'.⁶⁴ As a result, the state's exercise of power is legitimate so long as it is done to equally preserve, protect, and promote everyone's autonomy, even if this involves coercing the compliance of some people.

Even assuming that there is broad agreement on the need to enforce liberal institutions, considerable debate remains among liberals concerning what an equal concern for individual autonomy actually requires. Libertarian advocates of the 'minimal' state, for instance, adopt a narrow reading of the harm principle, claiming that people should be free to act as they please, provided only that they themselves do not actively harm others.⁶⁵ From this perspective, the state may only secure a space in which its constitutive members are free to develop and exercise their autonomy without the deliberate obstruction or interference of other people. It is therefore permissible, according to these libertarians, for the state to collect taxes in order to establish protective agencies such as the military, police, and criminal justice system. Raising taxes to fund other public institutions that do not directly provide for the common defense of individual liberty, however, is equated with unjustified coercion and oppression.⁶⁶

This minimal conception of the state is not popular with liberal egalitarians. They contend that libertarians presume an extremely impoverished notion of

Kant's *Critique of Practical Reason* (1788) [5:62–65] in Kant, *Practical Philosophy*, pp. 190–192. Also see 'On the Common Saying: That May Be Correct in Theory, but It Is of No Use in Practice' (1793). This theme is also taken up in Rawls, 'Basic Liberties and Their Priority' and 'Priority of Right and Ideas of the Good'.

⁶⁴*The Metaphysics of Morals* (1797) [6:232] in Kant, *Practical Philosophy*, p. 388 (Kant's emphasis).

⁶⁵Exemplars of this approach are Hayek, *The Constitution of Liberty*; Milton Friedman, *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962), Friedman and Friedman, *Free to Choose*; and Nozick, *Anarchy, State, and Utopia*.

⁶⁶According to *ibid.*, p. 169: 'Taxation of earnings from labor is on a par with forced labor'.

individual liberty and the harm principle.⁶⁷ In order to truly be free, according to liberal egalitarians, a person must have a real possibility for taking advantage of her freedom.⁶⁸ The state may promise a wide variety of liberties for everyone, but those without money, homes, jobs, good health, or an education do not seem as free to pursue their conception of the good life as those with all these advantages. This view suggests that poverty, poor health, and ignorance, for example, obstruct the development and exercise of a person's autonomy just as much as criminals and hostile foreign powers. As a result, liberal egalitarians argue that the state must have more than a defense policy: it must provide other services, such as social security, healthcare, public education, and infrastructure for transportation, garbage disposal, running water, electricity, and other utilities. This requires the state to raise taxes to fund these services. Even so, liberal egalitarians argue that doing so is justified because these public goods ensure that the freedom of each of the state's members has genuine value, which is necessary for genuinely respecting and promoting their autonomy.

Given the fact of reasonable pluralism, disagreements like these should not be surprising. As I noted in the previous section, agreement on adopting liberal institutions does not require agreement on a single comprehensive doctrine. In this case, libertarians and liberal egalitarians endorse competing comprehensive doctrines concerning the extent to which economic redistribution by the state is permissible. If each group takes the authoritative approach, then no matter what redistribution scheme the state adopts, at least one group is likely to believe that this scheme is—by this group's own lights—the incorrect scheme and that this therefore makes it illegitimate.⁶⁹ This illustrates a more general problem: it does not seem possible on this approach for a liberal to disagree with a particular exercise of state power and yet also regard it as legitimate. Without any uncontroversial resolutions to the conflicts dividing liberals, there remain open questions regarding the authoritative framework concerning the laws and decisions that the liberal state may legitimately enforce.

⁶⁷See, for instance, Bernard Manin, 'On Legitimacy and Political Deliberation', *Political Theory* 15, 3 (August 1987): 338–368, p. 339; Korsgaard, 'Commentary on G. A. Cohen and Amartya Sen', pp. 58–59; Anderson, 'What Is the Point of Equality?', pp. 314–315; and Martha C. Nussbaum, 'Forward', in *Libertarianism: For and Against*, ed. by Craig Duncan and Tibor R. Machan (Lanham, MD: Rowman and Littlefield, 2005): ix–xiv, p. xii.

⁶⁸As Rawls, *Theory of Justice*, pp. 204–205, 224–228, puts it, there is a concern not simply for liberty but for the 'worth of liberty' or its 'fair value', though Rawls takes these notions in a slightly different direction than I do here.

⁶⁹This is a problem connected with a more general class of 'correctness' theories of legitimacy in David M. Estlund, 'Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 173–204, p. 175.

This is a problem the consensual approach to political liberalism attempts to avoid. It claims that the commitment to liberty and equality prescribes certain tests for assessing the state's exercise of power in a given situation and that only those exercises satisfying the appropriate test are legitimate. In particular, this approach suggests that legitimacy requires that the state's constitutive members collectively authorize the use of state power. As a result, the consensual method attempts to sidestep the disagreements bedeviling the authoritative approach. The idea is that it may be possible for the state to obtain authorization to exercise its power in certain situations, even if it is an open question concerning what an equal concern for individual autonomy actually requires in those situations. The state may impose a taxation scheme, for instance, without waiting for libertarians and liberal egalitarians to settle their differences.

As with the authoritative approach to legitimacy, the consensual approach maintains that liberty is not license to be bound by no laws. Unlike the authoritative formulation, though, the emphasis is not on how respecting the autonomy of other people places limits or constraints on one's actions. Instead, it argues that respecting a person's autonomy involves only limiting or constraining her by laws that are, in some sense, of her own making. This is a classic attitude from the Enlightenment: a person is only free when she obeys laws that she has made and laid down for herself.⁷⁰ In this case, coercing her to conform to these laws does not disrespect her autonomy; such coercion rather affirms her autonomy and integrity by honoring her own decisions. In a similar fashion, if the state exercises its power only when it has the collective authorization of its constitutive members, it is then only enforcing compliance with laws and decisions its people have already made and laid down for themselves. If this test is not met and the state acts all the same, then it is unclear how the state is respecting its people's autonomy. It instead appears to be an instance of unjustified coercion or oppression.

The consensual structure regarding legitimacy requires that exercises of power by the state remain responsive to the claims and judgments of its constituent members. One immediate problem, however, with requiring such responsiveness by the state is whether it can account for why the state should adopt basic liberal institutions. It is extremely unlikely, to say the least, that neo-Nazis and religious

⁷⁰This is essentially Kant's argument from the *Groundwork of the Metaphysics of Morals* (1785), especially section II [4:430–445] in Kant, *Practical Philosophy*, pp. 81–93. The connection to the Enlightenment is made explicitly by Kant in 'An Answer to the Question: What is Enlightenment?' (1784). Kant is here drawing on the position that Rousseau defends in *On the Social Contract* (1762). For instance, see book I, chapters vi–vii and book II, chapters iii–iv, vi in Rousseau, *On the Social Contract*, pp. 52–55, 61–67. Rawls, *Political Liberalism*, p. 77 similarly maintains that 'full autonomy is realized by citizens when they act from principles of justice that specify the fair terms of cooperation they would give to themselves when fairly represented as free and equal persons.'

fundamentalists would ever willingly agree to restrain themselves from coercing others to embrace or otherwise conform to their particular conceptions of the good life. If they refuse to agree to authorize the state, the concern is whether state enforcement of basic liberties is now illegitimate. From the liberal vantage, it seems perverse to allow the overturning of individual liberties because certain intolerant groups do not see fit to recognize them.

In order for the consensual approach to respond to this and similar problems, a great deal hinges on how the collective authorization of state power is understood. There are two prominent positions on this.⁷¹ The first is the voluntaristic position. According to it, legitimacy requires that each and every one of the state's constitutive members actually agree to authorize the state to act.⁷² This does not seem encouraging, for intolerant groups simply refuse to accept liberal institutions. Even so, this position need not advocate that any presumptive agreement is authoritative. It can allow that agreements made under duress, misinformation, or ignorance, for example, are not binding. The problem remains that this does not obviously rule out the positions of all intolerant people. To the extent that it demands an actual agreement for authorizing the state's exercise of power, the voluntaristic position seems unable to sustain even the basic elements of the liberal state.⁷³

The second is the rationalistic position and it seems more promising. According to this view, legitimacy requires that the state's constitutive members, acting both rationally and reasonably, agree to authorize the state to act. On this view, people behave rationally insofar as they each authorize state action only when doing so has been presented with adequate justification, and they behave reasonably insofar as they each accept that such a justification requires reasons that they believe no reasonable person in their situation could deny. Bringing this together, the rationalistic position claims that the exercise of power by the state in a given situation is legitimate only when it is, at least in principle, justifiable in terms that cannot be reasonably rejected by persons committed to reach an agreement in that situation.⁷⁴

⁷¹I take this distinction from Waldron, 'Theoretical Foundations of Liberalism', p. 140.

⁷²Examples of liberal theories displaying elements of the voluntaristic position include Locke's *Second Treatise of Government* (1690); Rousseau's *On the Social Contract* (1762); and Simmons, *Moral Principles and Political Obligations*, 'Justification and Legitimacy'.

⁷³The voluntaristic approach might respond that liberal institutions are necessary prerequisites in order for the state's members to reach a free and informed agreement. Freedom of expression, for instance, might be necessary in order that everyone may be fully informed before coming into any agreement. This alone, however, seems to place liberal institutions on rather contingent ground.

⁷⁴I adopt this rather laborious phrase in light of Simon Căbulea May, 'Religious Democracy and the Liberal Principle of Legitimacy', *Philosophy and Public Affairs* 37, 2 (Spring 2009): 136–170, pp. 142–146. May argues that notion of reciprocity (i.e., that reasonable persons propose terms of

I stress that while the rationalistic position on legitimacy involves a certain amount of idealization concerning the behavior (they are rational and reasonable) and motivation (they are committed to reach an agreement) of the state's constitutive members, it need not entail a full-blown hypothetical approach to legitimacy.⁷⁵ That is, the legitimacy of *actual* state actions need not depend upon what an entirely *imaginary* group of people could agree upon. The hypothetical approach is certainly one way in which liberals may interpret the rationalistic position, but it is not the only way.⁷⁶ The rationalistic position is also perfectly consistent with claiming that the legitimacy of state action depends upon the actual agreements made between the state's actual members, provided that these people conform to the behavioral and motivational requirements. The extent to which real people can and will conform to these conditions is a separate matter.

Whether the people in involved are hypothetical or actual, the concern remains that the fact of reasonable pluralism renders the rationalistic position too strong. In light of the burdens of judgment from section 1.2.2, it is possible that in controversial situations every available choice for state action or inaction would be reasonably rejected.⁷⁷ If this is true, then there is no legitimate response, including inaction, by the state in that situation. One way to ensure that the rationalistic

cooperation that they sincerely believe are justified by reasons that other reasonable people may accept, see n. 40) must be made explicit in this condition of legitimacy by claiming that the persons are committed to reaching an agreement in their particular situation. I also follow Thomas M. Scanlon, 'Contractualism and Utilitarianism', in *Utilitarianism and Beyond*, ed. by Amartya K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982): 103–128, pp. 111–112, by using 'none could reasonably reject' in this formulation instead of 'all could reasonably accept' because the latter seems to admit too much. To illustrate, consider a state that uses its power to force all its Catholics—and no one else—to strictly adhere to Catholic orthodoxy. Suppose also that these Catholics all believe, in fact, that following their church's doctrine is absolutely necessary for their eternal salvation. It is plausible then that it would not be unreasonable for each of them to authorize this use of state power. Even so, it hardly seems unreasonable for one of them to reject this proposal by, for instance, making Locke's argument from *A Letter Concerning Toleration* (1689) that it would be inappropriate for her to abandon the care of her own salvation to the care of the state. On this, see volume 5 of Locke, *Works*, pp. 10–11.

⁷⁵I am drawing on the account of James S. Fishkin, 'Towards a New Social Contract', *Noûs* 24, 2 (April 1990): 217–226, pp. 217–218, which separates (1) the distinction between 'brute' and 'refined' motivations from (2) the distinction between 'actual' and 'hypothetical' choice situations. The distinction between voluntaristic and rationalistic positions may suggest a distinction between brute and actual agreement, on the one hand, and refined and hypothetical agreement, on the other. This certainly is the case in Waldron, 'Theoretical Foundations of Liberalism', p. 140. However, while my account here does treat the voluntaristic position as both brute and actual, I allow that the rationalistic position may be either actual or hypothetical.

⁷⁶I address both hypothetical and actual choice situations respectively in sections 2.3 and 2.4 of chapter 2.

⁷⁷For instance, see David Reidy, 'Reciprocity and Reasonable Disagreement: From Liberal to Democratic Legitimacy', *Philosophical Studies* 132, 2 (January 2007): 243–291, pp. 265–272.

position is not so strong is to emphasize that the involved parties are committed to reaching an agreement of some kind in their particular situation.⁷⁸ Just because reasonable disagreement is possible does not mean that someone must reasonably disagree. A person's motivation to reach an agreement with others might overrule her desire to reasonably disagree in certain situations.

A second way to address this concern is to adjust the scope of application for the rationalistic position. Rather than demand that the state's constitutive members authorize every single exercise of power, the rationalistic account may claim that these people must only agree on certain constitutional essentials to which exercises of power by the state must adhere.⁷⁹ By constitutional essentials, I mean the fundamental principles prescribing, first, when and how the state may exercise its power in particular situations and, second, the basic liberties of the state's constitutive members that the state must respect and promote.⁸⁰ The assumption is that agreement on these fundamental principles is more likely than agreement on every single exercise of state power, and that these principles in turn dictate how the state should act when reasonable conflicts and disagreements concerning divide its constitutive members in particular situations. Legitimacy is therefore derivative from an agreement on constitutional essentials. The rationalistic position may then hold that the exercise of power by the state in a given situation is legitimate only when done in accordance with a constitution whose terms cannot, at least in principle, be reasonably rejected by persons committed to reach an agreement in that situation. I call this the liberal test of legitimacy.

I believe the dominant programmatic task in contemporary liberalism is to construct a rationalistic consensual approach to legitimacy along the lines of this test.⁸¹ This program argues that basic liberal institutions should be among the

⁷⁸This is the response of May, 'Religious Democracy and the Liberal Principle of Legitimacy', pp. 144–146.

⁷⁹This distinction between agreement on constitutional essentials versus agreement on 'ordinary-level political acts' is put nicely by Frank I. Michelman, 'The Subject of Liberalism', *Stanford Law Review* 46, 6 (July 1994): 1807–1833, pp. 1827–1828, 'Review of Jürgen Habermas' *Between Facts and Norms*', *Journal of Philosophy* 93, 6 (June 1996): 307–315, pp. 309–310 and 'How Can the People Ever Make the Laws? A Critique of Deliberative Democracy', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 145–171, p. 154.

⁸⁰For instance, see Rawls, *Political Liberalism*, p. 227.

⁸¹Variations on this approach are found in, among many others, Ackerman, *Social Justice in the Liberal State*, pp. 4–8; Joshua Cohen, 'Deliberation and Democratic Legitimacy', in *The Good Polity: Normative Analysis of the State*, ed. by Alan Hamlin and Philip Pettit (Oxford: Basil Blackwell, 1989): 17–34, p. 22; Young, *Justice and the Politics of Difference*, p. 34; Nagel, *Equality and Partiality*, p. 33; Rawls, *Political Liberalism*, p. 137; Amy Gutmann and Dennis Thompson, *Democracy and Disagreement* (Cambridge, MA: Belknap Press of Harvard University Press, 1996), p. 99; Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*,

state's constitutional essentials by suggesting that those who reject them are not being reasonable. Given the account of reasonableness from section 1.2.2, this tactic may be plausible. This program also argues that it is possible for someone holding a reasonable comprehensive doctrine to disagree with the exercise of state power in a particular situation and yet not be able to impugn its legitimacy; disagreement does not immediately imply illegitimacy, as it does with the authoritative approach. On this score, libertarians and liberal egalitarians, for instance, may not have to resolve their differences in order for the state to enforce some specific scheme of economic redistribution, provided that it is derived from constitutional essentials not open to reasonable rejection by either of them.⁸² Similarly, just because an outcome does not further one's own conception of the good life need not imply that one can reasonably reject its basis; simply not liking an outcome is not enough.

The specifics of how the liberal test of legitimacy works varies from theory to theory, but this at least gives a sense of the rationalistic consensual paradigm. The issue that concerns me is that this test potentially creates significant vagaries as to how the state should make decisions when several courses of action pass the liberal test of legitimacy. I believe this concern becomes manifest when liberals conclude that this test along with the fact of reasonable pluralism entail that the liberal state should be neutral when making its laws and decisions. In the next section, I argue that the liberal test of legitimacy, insofar as it demands state neutrality, comes into tension with the need for the state to possess a consistent and coherent political agenda.

trans. by William Rehg (Cambridge, MA: MIT Press, 1996), p. 107, Michelman, 'How Can the People Ever Make the Laws? A Critique of Deliberative Democracy', pp. 153–154; Thomas M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Belknap Press of Harvard University Press, 1998), p. 153; and Estlund, *Democratic Authority*, p. 41. The rationalistic consensual approach to legitimacy is nearly ubiquitous in deliberative democracy, which is a popular movement amongst liberals. According to this view, the outcomes of public deliberation amongst the state's constitutive members, when suitably organized, reveals what passes the liberal test of legitimacy while simultaneously conveying to these people why this is so. Put more succinctly, deliberation is thought to ensure both legitimacy and stability. For an overview of this deliberative approach to liberalism, see Samuel Freeman, 'Deliberative Democracy: A Sympathetic Comment', *Philosophy and Public Affairs* 29, 4 (Autumn 2000): 371–418; and Amy Gutmann and Dennis Thompson, 'What Deliberative Democracy Means', in *Why Deliberative Democracy?* (Princeton: Princeton University Press, 2004): 1–63.

⁸²Michael J. Sandel, 'Political Liberalism', *Harvard Law Review* 107, 7 (May 1994): 1765–1794, pp. 1784–1786, argues that Rawls implicitly assumes that the libertarian position is not reasonable because it rejects Rawls' 'difference principle'. According to Sandel, Rawls apparently believes there is no reasonable way to reject this principle. John Rawls, *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999), p. 49, seems to confirm this. This does run the risk of making the concept of reasonableness both arbitrary and exclusive. As far as I am concerned, however, both libertarianism and liberal egalitarianism are reasonable doctrines.

1.4 Responsive Impartiality and State Agency

Liberalism's commitment to liberty and equality requires that the liberal state secure a robust sphere of individual liberty for all of its constitutive members. Doing this permits, even if it does not actively encourage, a reasonable pluralism of comprehensive doctrines. As a result, the liberal state must anticipate what to do when its people inevitably endorse competing doctrines that make conflicting claims on the behavior of the state's members, on the social and material resources of the state, and on the very structure of the state itself. Debates between reasonable people not only involve long-standing controversies such as abortion, homosexuality, and torture, but also more routine issues such as law enforcement, national defense, market regulation, social security, healthcare, education, environmental protection, and the distribution of financial burdens. In these circumstances, the state cannot simply hope to settle things by asking the parties to 'be reasonable'.

The dominant approach in contemporary liberalism claims that liberty and equality require that no matter how the state responds to societal conflicts, its exercises of power must be in accord with a constitution whose terms its constitutive members cannot reasonably reject. Bringing the fact of reasonable pluralism to bear on this demand of legitimacy suggests to many liberals that the liberal state must remain neutral—or as I prefer to say, impartial—regarding the reasonable comprehensive doctrines that its constitutive members may adopt if the state is to make legitimately binding decisions.⁸³ Impartiality in this context has two aspects. The first comes from the demand that the liberal state tolerate the acceptance of any reasonable comprehensive doctrines by its constitutive members. Liberal proponents of impartiality typically require that the state go further by ensuring that each of its constitutive members has an equal opportunity to pursue the reasonable doctrine of her choice.⁸⁴ The second aspect of impartiality prohibits the state from exercising its power in order to promote or otherwise favor any particular doctrine. This is not to say that each and every exercise of state power must affect the adherents of all reasonable doctrines equally; this may be impossible to satisfy, as any non-trivial use of state power almost invariably benefits some and not others. The claim is rather that the state may only justify its actions by appealing to reasons and principles—either codified in the constitution or consist-

⁸³Some representative contemporary proponents of state neutrality include Rawls, *Theory of Justice, Political Liberalism*; Nozick, *Anarchy, State, and Utopia*; Dworkin, 'Liberalism', 'Why Liberals Should Believe in Equality', *New York Review of Books* 30, 1 (3rd February 1983): 32–34; Ackerman, *Social Justice in the Liberal State*, 'Neuralities', in *Liberalism and the Good* (London: Routledge, 1990): 29–43; Charles E. Larmore, *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987); Will Kymlicka, *Liberalism, Community, and Culture* (Oxford: Oxford University Press, 1989); and Nagel, *Equality and Partiality*.

⁸⁴See, for instance, Rawls, 'Priority of Right and Ideas of the Good', pp. 260–264.

ent with those so codified—that none of its constitutive members can reasonably reject. The state must therefore be able to justify its actions by appealing to reasons that are themselves impartial between the reasonable comprehensive doctrines its members endorse.⁸⁵ Both these aspects of impartiality suggest, at the very least, that the state has no business in taking sides in any conflicts or disagreements that might emerge between reasonable doctrines.

Liberal defenders of impartiality typically argue that it is necessary in order for the state to equally respect the autonomy of all its constitutive members.⁸⁶ The claim is that the state does not show equal respect when it justifies an exercise of power by appealing to principles grounded by a doctrine that some of its constitutive members reasonably reject. Doing so is thought to inform certain reasonable people that their judgments on political issues do not matter and demand that these people conform to a law or decision that, by their own lights, has no good reasons justifying its enforcement by the state. Nor is it equally respectful for the state, insist these liberals, to provide public resources or protection to those pursuing certain conceptions of the good life while withholding assistance from or enforcing punishments against those who hold other reasonable views.⁸⁷ It is especially impermissible on this view for the state to manipulate or coerce a reasonable person into regarding her favored way of life as having less value or meriting less concern than those adopted by others.

For such reasons, many liberals regard impartiality as a necessary consequence of the liberal test of legitimacy given the fact of reasonable pluralism. According to these liberals, coercion is legitimate when done to secure a space for the state's

⁸⁵For more on the distinction between impartiality of outcomes and impartiality of justifications, see Ackerman, *Social Justice in the Liberal State*, p. 73; Larmore, *Patterns of Moral Complexity*, p. 44; Rawls, 'Priority of Right and Ideas of the Good', pp. 260–264; Jeremy Waldron, 'Legislation and Moral Neutrality', in *Liberal Neutrality*, ed. by Robert E. Goodin and Andrew Reeve (London: Routledge, 1989): 61–83, pp. 66–68; Will Kymlicka, 'Liberal Individualism and Liberal Neutrality', *Ethics* 99, 4 (July 1989): 883–905, pp. 883–886; and George Sher, *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997), pp. 4, 22–23. This distinction is rejected, however, by Robert E. Goodin and Andrew Reeve, 'Do Neutral Institutions Add Up to a Neutral State?', in *Liberal Neutrality*, ed. by Robert E. Goodin and Andrew Reeve (London: Routledge, 1989): 193–210, p. 202. Goodin and Reeve argue that the outcomes or effects of state action, and not simply the reasons behind it, should also be neutral with respect to reasonable comprehensive doctrines.

⁸⁶This is the argument in Dworkin, 'Why Liberals Should Believe in Equality'.

⁸⁷Sher, *Beyond Neutrality*, pp. 34–37, suggests that using punishments and rewards are but two ways that the state might show favor for certain doctrines over others. The state might also adopt 'nonrational manipulation', say through certain forms of education, or selectively create background social circumstances, say in only recognizing certain sorts of marriages, in order to cause its constitutive members to accept certain doctrines rather than others. I focus on rewards and punishments because they are usually the most egregious and identifiable forms of coercion and manipulation that the state may adopt.

constitutive members to equally develop and exercise their individual autonomy; coercing a person is illegitimate when it is justified by appeal to the principles of a comprehensive doctrine that she reasonably rejects. Requiring state neutrality is supposed to ensure that the liberal state does not exceed this limit.

A tension begins to emerge, though, when one considers that as a critical locus of coordination amongst its members, the liberal state ought to adopt and act upon a consistent and coherent political agenda, one contained within or supported by its constitution. That is, the state must specify which issues count as matters of state concern and which do not. On matters of state, it must then make and enforce its decisions regarding them. In doing so, it seems contrary to liberalism—especially its roots in the European Enlightenment—for the state to make decisions in an arbitrary fashion or act in an ad hoc manner. Instead, the state should possess an overarching rationale, in accord with the liberal test of legitimacy, accounting for its particular agenda and for how it intends to pursue this agenda. A consistent and coherent political agenda is therefore derived from prior agreements on the fundamental principles of the state's constitution or from ongoing legislative reforms, and this agenda in turn grounds state agency by specifying the services the state offers, at what cost, and to whom. This leads to the worry, however, that the state must adopt a comprehensive doctrine of its own in order to provide the overarching rationale for its actions. Recall from section 1.2 that it seems that a systematic doctrine is necessary in order to avoid making ad hoc and arbitrary decisions. Consequently, as the state tries to make decisions that can be justified as not arbitrarily chosen, the concern is that the state is forced into adopting an increasingly systematic doctrine. If this is the case, the liberal state must apparently give up pretensions to impartiality. State agency seems on a collision course with state neutrality.

The need for the liberal state to adopt its own reasonable and systematic doctrine becomes more pronounced when considering the difficulties that reasonable pluralism poses for pursuing a consistent and coherent political agenda. Presumably, liberals believe that people cannot reasonably reject certain constitutional essentials, such as those protecting and promoting basic liberal institutions. Even so, the fundamental principles of a liberal constitution must be quite general if they are to be the object of agreement across a wide range of reasonable doctrines. The problem is that when reasonable debates divide the liberal state's constitutive members concerning whether a particular issue belongs on the political agenda or how the state should act on the issues already on its agenda, applying general principles may prove impossible because they are too abstract for direct application. In addition, attempting to specify principles that do apply seems likely to be a matter of further reasonable disagreement. For instance, even those who maintain that the liberal state should only ensure there is an open market that allows people to

forge their own voluntary, private agreements concerning these services reasonably disagree over what exactly constitutes an 'open' market and what rules the state may enforce to prevent market inefficiencies. The assurance of basic liberal institutions does not settle this matter. Similar concerns exist for social security, healthcare, environmental protection, and a whole host of other political issues. With reasonable disagreements practically unavoidable when constructing and acting on a political agenda, it is difficult to imagine an impartial overarching rationale on which reasonable people could agree.

It may therefore be tempting to respond to pluralism by having the liberal state act as a third party distinct from that of its constituent members. Under this arrangement, the liberal state occupies an independent standpoint from which to assess and order the competing claims its members may make concerning its political agenda. In doing so, the state adopts a constitution that codifies a systematic and reasonable doctrine of its own, which the state uses for setting more or less determinate weights over the competing values and principles in order to set and pursue a political agenda. This need not preclude the possibility of revising this doctrine over time through some amendment process as it is called upon to resolve new social controversies. This is just an increase of the doctrine's systematicity. Even so, at any given time, the state has its own doctrine, which is added to the extant societal pluralism, and the state declares that this doctrine alone trumps all others on matters of state. Matters of state aside, individuals are free to pursue their own reasonable doctrines. The state's doctrine therefore ensures that its agenda is coherent and consistent and thus social order is maintained. In this case, though, the state hardly seems neutral. The fact of reasonable pluralism holds that no social consensus exists on a reasonable doctrine, and so whatever doctrine the state adopts, it is possible that some of the state's constitutive members may reasonably disagree with it. With a state-sanctioned doctrine independent from the doctrines of its reasonable members, this situation begins to resemble a form of dictatorship. Consequently, I believe the form of state neutrality that liberals advocate is best captured by the term 'responsive impartiality'. The state ought to remain impartial for the reasons given at the beginning of this section, but it also ought to remain responsive to the positions and doctrines adopted by its constitutive members rather than dictatorially imposing its own. In section 3.6, I provide more substantive content to what I believe it means for the state to be responsively impartial; I believe at this point the meaning is sufficiently clear.

While responsive impartiality appears to ensure a commitment to liberty and equality that dictatorships lack, responsive impartiality also seems to confound the possibility for pursuing a consistent and coherent political agenda. On the one hand, it is not clear what fundamental principles for regulating the exercise of state power are both responsive to and impartial between the competing reasonable

doctrines adopted by the state's constitutive members. Responsive impartiality does not seem to reveal which position the liberal state should take should reasonable debates emerge on controversial issues such as social security or healthcare. If this is the case, responsive impartiality alone offers little guidance and adhering to it mires the liberal state in reasonable pluralism without a consistent and coherent course of action. So responsive impartiality is supposed to disentangle the liberal state from the various comprehensive views of its people, but it simply renders the state impotent.⁸⁸

On the other hand, if the liberal state acts on a consistent and coherent political agenda, it seems that either this reveals the systematic doctrine to which the state already subscribes, or this entails that the state begin to adopt a more and more systematic doctrine. In either case, the liberal state is not really impartial at all. In fact, one common complaint about liberal neutrality is that it is cover for liberals to advance their own comprehensive views. For instance, some liberals argue that neutrality requires that the liberal state tolerate controversial practices such as the Nazis march through Skokie, pornography, abortion, and same-sex marriage.⁸⁹ If the state restricts these, the argument goes, it essentially asks reasonable people adhere to the demands of certain doctrines that they may reasonably not accept, thus not respecting their autonomy and thereby violating neutrality. The suspicion is that the default 'impartial' position on these controversial issues turns out to be the conventional 'liberal' position, apparently demonstrating that liberal neutrality

⁸⁸A similar problem concerning the recent 'politics of difference' is recognized by Sheldon S. Wolin, 'Democracy, Difference, and Re-Cognition', *Political Theory* 21, 3 (August 1993): 464–483, p. 480:

There is, however, a political paradox in the re-cognition of difference. I want to be bound only by a weak and attenuated bond of inclusion, yet my demands presuppose a strong State, one capable of protecting me in an increasingly racist and violent society and assisting me amidst increasingly uncertain economic prospects. A society with a multitude of organized, vigorous, and self-conscious differences produces not a strong State but an erratic one that is capable of reckless military adventures abroad and partisan, arbitrary actions at home—oscillating as it were between Watergate and Desert Storm—yet is reduced to impotence when attempting to remedy structural injustices or to engage in long-range planning in matters such as education, environmental protection, racial relations, and economic strategies.

The context here concerns the situation of a person who wants to preserve her own personal cultural heritage with minimal inclusion in the political community, and yet the state cannot function if each of its members has this attitude with respect to her respective culture. The tension that I am describing concerns the demand that the state 'preserve' the comprehensive doctrine of each of its members by remaining impartial between them, and the problem the state has in taking effective action on controversial issues while doing so.

⁸⁹See especially Dworkin, 'Is There a Right to Pornography?' and 'Three Questions for America', *New York Review of Books* 53, 14 (21st September 2006): 24–30.

is not so neutral after all.

A related, though deeper, charge against state neutrality concerns why it, along with the commitment to freedom and equality, outweighs other important values.⁹⁰ Liberals usually appreciate that issues like Nazi demonstrations, pornography, and abortion may involve the dignity and welfare of the people affected by them, and these seem precisely the types of concerns that the liberal state should take some interest in addressing. If responsive impartiality prohibits the liberal state from acting on these issues, then liberal neutrality is again not neutral, but actively prejudiced against all competing considerations. These all point to the idea that it is contradictory for the liberal state to act definitively on important social issues while remaining responsively impartial; impartiality and effective state agency appear incompatible.

This is not meant to imply that further inquiry might reveal genuinely impartial principles that the state may use to construct a consistent and coherent political agenda. Rather, the concern is that the state must make decisions now and cannot wait until inquiry is completed or a consensus over principles emerges. In acting on its own agenda, it seems that the state faces the problem of setting more or less determinate weights over competing values and principles. In other words, state agency apparently requires that the state adopt a doctrine with an ever increasing measure of systematicity in order to resolve societal conflicts. Without a social consensus, however, the worry is that there remain those who reasonably disagree with whatever agenda or weighing of social commitments and values the state adopts. The state cannot avoid taking controversial positions on certain issues that appear dictatorial to those who reasonably disagree with those positions. Since the state remains accountable to these people, the challenge is understanding how the state may take legitimate action on issues that admit of reasonable disagreements and conflicts.

The dilemma facing the liberal test of legitimacy ultimately comes down to this: if political legitimacy requires that the liberal state remain responsively impartial between the competing comprehensive doctrines of its constitutive members, then the state does not seem able to make principled decisions. And if the state seeks to make principled decisions, it seems necessary for it to adopt and then impose its own reasonable and systematic doctrine upon its constituent members and therefore run the risk of exercising its power illegitimately. The state is either

⁹⁰This is the common charge made by communitarian critics of liberalism like Charles Taylor, 'What's Wrong with Negative Liberty?', in *The Idea of Freedom: Essays in Honour of Isaiah Berlin*, ed. by Alan Ryan (Oxford: Oxford University Press, 1979): 175–193; Michael J. Sandel, *Liberalism and the Limits of Justice*, 2nd ed. (Cambridge: Cambridge University Press, 1998); and Alasdair MacIntyre, *After Virtue: A Study in Moral Theory*, 3rd ed. (Notre Dame: University of Notre Dame Press, 2007).

rendered impotent by the fact of reasonable pluralism or dictatorially imposing its own systematic doctrine. The underlying conclusion of this chapter is worth stressing once more: in order to avoid making ad hoc and arbitrary decisions, the concern is that the state must adopt more and more systematic doctrines, and it is this systematicity that comes into tension with responsive impartiality in the face of societal pluralism. As a result, liberalism must navigate this dilemma with an account that allows the state to remain responsively impartial with respect to reasonable disagreements between its constitutive members while also pursuing a consistent and coherent political agenda of its own.

While I believe this dilemma is serious, for it lies behind everything that follows in this dissertation, it is thought to have an obvious solution: the liberal state must employ a system of democratic electoral and legislative processes. It is such a system of democratic institutions that allow the liberal state to coherently and consistently construct a social agenda and act upon it. While I certainly agree that liberals ought to embrace democracy, I do not believe that doing so is sufficient to evade this dilemma. In particular, the results of social choice theory—I have in mind Kenneth Arrow's famous theorem among others—demonstrate how this dilemma remains even when democratic institutions are in place. Assessing this theorem, though, also reveals what I take to be the proper response to the tension between responsive impartiality and state agency: rejecting the very idea that a systematic view is necessary for the state to adopt a consistent and coherent social agenda. I get ahead of myself, however. Before going there, I must first outline the various ways in which liberal theorists have typically presumed that democratic processes ease that tension. I now turn to this.

CHAPTER TWO

Liberal Democratic Proceduralism

ACCORDING to the rationalistic consensual approach to legitimacy, exercises of power by the state are legitimate only when they are justifiable in terms that cannot be reasonably rejected by persons committed to make a decision in that situation. The fact of reasonable pluralism suggests that this approach requires that the state remain neutral between the different and often competing comprehensive doctrines that inevitably emerge in a liberal state. This demand comes in tension with state agency, for it seems difficult for the liberal state to pursue a consistent and coherent social agenda when doing so appears to require the state to take a stand on controversial social issues. State agency therefore apparently requires that the state possess its own reasonable comprehensive doctrine to give order and coherence to its actions—but then the liberal state appears to impose its views on those who may reasonably disagree with them.

Nevertheless, this dilemma is thought to have an obvious solution: a system of democratic institutions governing electoral and legislative processes that are open to all the constitutive members of the state. The reason is that a system of democratic processes including public deliberation, elections, referenda, legislative and executive orders, and collective bargaining is thought to offer neutral and impartial ways for the state to make decisions in the face of reasonable disagreements between its members. There are several reasons to believe this. First, democratic processes do not seem to entail a particular reasonable comprehensive doctrine or conception of the good life. Second, they are impartial between these doctrines and conceptions insofar as they do not, in and of themselves, express positive or negative judgments concerning these doctrines and conceptions. Third, they

offer mechanisms giving the state's constitutive members a chance to present their claims and to be heard. Fourth, they give each member an equal opportunity—one vote or a set time to present one's case—to influence the outcomes of the process.¹ Fifth, given the previous considerations, democratic processes do not impose a particular comprehensive doctrine on their participants, but they are instead responsive to the contents of the participants' doctrines. Sixth, this responsiveness is governed by established rules, rendering it consistent by treating similar cases in a similar manner and coherent by making it possible to anticipate its outcomes in delineated circumstances.

The last of these considerations reveals how democratic institutions may ease tensions between neutrality and state agency. Such institutions embody a form of procedural rule-governed neutrality that permits the state to act on political issues even while reasonable disagreements and significant political controversies may render consensus on them extremely unlikely. Private individuals must decide for themselves where they stand on the issues and they are all free to engage with others in order to convince them as to the merits of their respective positions. At their best, democratic institutions facilitate these interactions and allow everyone to be heard and influence the final decision, however it is ultimately made. Employing these processes therefore provides an underlying consistency and coherence of the state's agenda without requiring the state to impose its own comprehensive doctrine. Neutrality and state agency thus appear reconciled.

While democratic institutions appeal to contemporary liberals, this does not necessarily mean that liberals endorse a strictly majoritarian form of democracy where there are no limitations on the laws and decisions that a bare social or legislative majority may enact and enforce. Liberalism, after all, holds that basic liberal institutions cannot be abridged simply because a bare majority votes to overturn them. Unrestrained majority rule may disregard the equal moral worth of persons, for instance, by enslaving or otherwise oppressing minorities in order to benefit the majority. As a result, contemporary liberals typically advocate liberal, or constitutional, democracy, which adopts a constitution setting limits on what social and legislative majorities may accomplish. The exemplar here is usually the United States Constitution, which guarantees certain basic liberties

¹Certainly this and the previous claim may require more than making democratic institutions equally open to all of the state's constitutive members. Along lines similar to those followed by liberal egalitarians in section 1.3, liberals may worry that the process is open to everyone, but those without money, homes, jobs, good health, or an education may not be able to engage in the process on equal terms as those with all these goods. Those with more education may know how to use these processes to their benefit while those with more money might be able to create more time to devote to these processes. This puts those in extreme poverty at an extreme disadvantage. Making democratic processes treat everyone as equals may then require significant background social conditions that extend well beyond the processes themselves.

in a bill of rights and authorizes a supreme court to annul the decisions of bare majorities should they violate those liberties. Additional checks on majority rule are the separation of powers, federalism, bicameral legislatures, and a stringent constitutional amendment process requiring supermajorities at various levels of government. This is, however, only one way to organize a liberal democracy; different societies may find alternative arrangements that protect and promote basic liberal institutions.

There is also second, related sense in which contemporary liberals understand a liberal democracy as opposed to a majoritarian democracy. In a majoritarian democracy, state decisions need only remain responsive to the interests, whatever their content, of the current social or legislative majority. The interests of minorities are promoted only to the extent that they overlap with those of the majority. A liberal democracy, on the other hand, is supposed to promote the common good of everyone by making decisions that mutually benefit all of the state's constitutive members. According to liberalism, everyone shares an interest in freely developing and exercising their autonomy and so a liberal democracy promotes this common good by protecting and promoting basic liberal institutions. Beyond this, most contemporary liberals believe that the fact of reasonable pluralism obscures further prescriptions for the common good, leaving it to the institutions and processes of a liberal democracy to work out the details of the common good of its particular members and to apply them when making laws and decisions.

The forgoing has only been intended as a sketch of the appeal to contemporary liberals for having the state govern with a system of democratic institutions and how these liberals understand the relationship between democratic and liberal institutions. The particulars vary greatly across liberal views, and my critique of liberal democratic proceduralism in the next chapter does not affect all these approaches equally, though it certainly challenges all of them. In this chapter, I present a taxonomy of various approaches to liberal democratic proceduralism. To begin, I provide a general orientation in section 2.1 by adapting John Rawls' categories of procedural justice to represent alternative ways in which a given procedure may be justified. I then proceed in sections 2.2–2.4 to discuss how these lead to different approaches for justifying democratic processes. In doing so, I follow a natural trajectory in contemporary liberalism—fueled by the fact of reasonable pluralism—that goes from emphasizing the quality of the outcomes that these processes produce to emphasizing the fairness of these processes independent of the outcomes they may produce.

2.1 Procedural Justifications

The fact of reasonable pluralism leads many liberals to suggest that there is little hope of consensus on a set of political standards that the state could appeal to for justifying its decisions on a broad range of issues.² This has caused dissatisfaction with the classical liberalisms of John Locke, Immanuel Kant, and John Stuart Mill, whose approaches are thought to rely on controversial comprehensive doctrines. Drawing on a common distinction from political theory, I call an approach to decision making by the state that emphasizes direct application of a set of political standards a substantive approach. The default position amongst many contemporary liberals appears to be that the fact of reasonable pluralism precludes such an approach.

In light of this position, many liberals increasingly emphasize the need for a system of fair procedures, such as deliberation and voting, that the state should employ when making decisions or organizing collective action. In contrast to substantive approaches, these are procedural approaches. In *A Theory of Justice*, Rawls characterizes aspects of his own position in terms of procedural justice.³ While Rawls' particular applications do not concern me yet, his taxonomy of procedural approaches suggests different tactics for justifying the use of procedures in a given situation. This taxonomy distinguishes between perfect and imperfect procedural justifications, both of which, in turn, are distinguished from pure procedural justifications.

Both perfect and imperfect procedural approaches justify a procedure by arguing that it ensures better outcomes according to some process-independent standard. The distinction between these concerns the feasibility of the procedure and how reliably it attains those desired outcomes. A perfect procedural justification argues that a given procedure is both feasible and guaranteed to return the right outcome. An imperfect procedural justification argues that the procedure is more or less feasible and that it is more or less reliable in attaining that outcome, though without any absolute guarantees. Perfect and imperfect procedural justifications therefore have three components: the specification of (1) what the procedure involves, (2) the procedure-independent standard specifying which outcomes are right, and (3) the background circumstances necessary to ensure the

²Stuart Hampshire, 'Liberalism: The New Twist', *New York Review of Books* 40, 14 (12th August 1993): 43–47, pp. 46–47 and Norman Daniels and James Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?* (Oxford: Oxford University Press, 2001), pp. 3–4 are the most explicit about this.

³John Rawls, *A Theory of Justice* (Cambridge, MA: Belknap Press of Harvard University Press, 1971), pp. 83–90, 120, 136, 304. In particular, he applies procedural notions to justify his two principles of justice resulting from the original position and to support the claim that the distribution of goods and services, whatever it may be, is just within a society satisfying these two principles.

feasibility of the procedure and its reliability for obtaining those outcomes.

Rawls uses the case of fairly dividing a cake between two people to illustrate a perfect procedural approach. If fairness calls for equal shares, then there is a procedure guaranteed to give this result: the first person cuts the cake and the second person decides who gets which piece.⁴ The reliability of this procedure, of course, depends on both people solely desiring to get as much cake as possible and this fact being known to them. A situation that Rawls gives for a more imperfect procedural approach is a criminal trial. Justice demands that a person is found guilty if and only if that person is, in fact, guilty. As of yet, though, there is no organization of the criminal justice system guaranteeing this. Nevertheless, certain conditions, such as following certain rules of evidence and the presentation of witnesses, are supposed to increase the reliability of criminal trials. In each of these instances, the procedure is justified insofar as it reliably produces the requisite outcomes.

As opposed to perfect and imperfect procedural justifications, pure procedural justifications do not appeal to the types of outcomes that the procedure produces but to features of the procedure itself. It is not enough, for instance, that the criminal justice system correctly ascertains innocence or guilt outcome: it must also treat the parties in certain ways. A piece of evidence may rightfully demonstrate the incontrovertible guilt of the accused, but if acquired in an inappropriate way—say without a search warrant—it is not admissible. Following these procedural elements is supposed to make for *fair* criminal proceedings, even though they may obstruct the process from returning the correct assignment of guilt or innocence. The general idea behind pure procedural justifications is that ‘there is a correct or fair procedure [for the given situation] such that the outcome is likewise correct or fair, whatever it is, provided that the procedure has been properly followed. . . . A fair procedure translates its fairness to the outcome.’⁵ In such a case, the fairness of an outcome is derivative from the fairness of the procedure producing it, not the other way around as in perfect and imperfect procedural justifications.⁶ Of course, the obvious question is what then makes a procedure fair if not its outcomes. One might worry that the decisions of the criminal justice system suffer by requiring search warrants. This seems to create a tension between the claim that there is no independent standard for evaluating outcomes and the assumption that not just

⁴An analogous procedure can be specified for fairly divided a cake between any given number of people. See Hugo Steinhaus, ‘The Problem of Fair Division’, *Econometrica* 16, 1 (January 1948): 101–104.

⁵Rawls, *Theory of Justice*, p. 86.

⁶It may be tempting to claim that in a pure procedural justification, the outcome is correct or fair only by appeal to some set of ‘intrinsic’ features of the process. I avoid this loaded term, however, because its connotations are far too broad and perhaps even inappropriate in this context.

any outcome is acceptable.

This tension need not be so serious, as Rawls' example of fair betting illustrates. Suppose a bookie offers Jones to flip a fair coin: if it lands heads up then Jones wins five dollars, and if it lands tails up she loses five dollars. Unlike the case of the criminal justice system, there is no procedure-independent standard for assessing the outcomes of this process. Without actually carrying out the gamble, there is no correct way to fairly distributing their money. This does not imply that any distribution is fair; the fair distribution is the one resulting from actually flipping the coin. In the case of betting, Rawls claims that there are relevant background circumstances that apply when assessing the fairness of the process: 'that fair bets are those having zero expectation of gain, that the bets are made voluntarily, that no one cheats, and so on'.⁷ When these circumstances are satisfied, the result of acting carrying out the procedure, no what distribution it entails, is fair.

Similar norms hold for virtually any organized sport. In professional golf, an assortment of rules governs the play of the game and how the competitors' scores are determined, and these may vary from tournament to tournament depending on whether they adopt the stroke play, match play, or Stableford scoring system. This means that there are no process-independent standards governing who should win a given golf tournament and gain the resulting accolades. The only way to determine the winner is to actually play.⁸ In this case, the background circumstances for a fair professional golf tournament are entirely conventional, having been selected by the Professional Golf Association and coifed in their rules books. As long as all the players follow the rules, the outcome of the game is fair.

Background circumstances form the set of motivating desiderata or criteria for setting the parameters within which a correct or fair procedure may freely operate. Without them, pure procedural justifications would be impossible: there would be no sensible way to distinguish between better and worse procedures for a given situation. It is only by appeal to the background circumstances of fair gambling that someone could meaningfully argue that a given bet was unfair because it had a high expectation of loss. Similarly in professional golf, an outcome is only unfair when it can be shown that a player violated the rules. Even a person rightfully found guilty in a court of law can appeal this decision provided that he can demonstrate that the decisive evidence against her was obtained improperly. So while a perfect or imperfect procedural justification appeals to the quality of the

⁷Rawls, *Theory of Justice*, p. 86.

⁸At times, people may label someone a good golfer by pointing to certain of her attributes, such as general endurance, manual dexterity, good eyesight, and accuracy in shotmaking. Even so, if she has not consistently done well when playing *actual* rounds of golf, calling her a good golfer is, at best, speculation or, at worst, erroneous. There are no standards for being a good golfer that are independent of actually playing the game well.

outcomes a procedure produces, a pure procedural one must look to the relevant background circumstances within which it is supposed to operate.

I emphasize the importance of the background circumstances for two reasons. In the first place, different situations seem to raise radically different background circumstances and so call for different procedures. The relevant circumstances for fair gambling are quite different than those for professional golf, and both of these seem to differ greatly from those that may apply to fair decision making by the state in the face of social disagreement. Even situations of social disagreement may differ substantively, as evinced by the wide variety of decision-making procedures that the state may employ: the state may facilitate bargaining between conflicting parties, it may have a recognized authority or mediator make the decision, or it may organize a vote. The relevant background circumstances seem to vary from case to case. For instance, the circumstances for fair bargaining may simply prohibit deception and threats of physical force, but not much else; those for fair mediation may require that each side get its due hearing and that the verdict is impartial; and those for fair voting may demand equal opportunities for participation and consideration for all voters. Reducing background circumstances to a general set of conditions for all pure procedural justifications does not seem promising.⁹

A second reason for emphasizing background circumstances is because background circumstances may have more than a justificatory role because they exert influence on the outcome of procedures conforming to them.¹⁰ Returning back to

⁹I mention this because William Nelson, 'The Very Idea of Pure Procedural Justice', *Ethics* 90, 4 (July 1980): 502–511 argues that pure procedural justifications may only appeal to the entitlements of those participating in the procedure and the free exercise of those entitlements. According to Nelson, the only relevant background circumstances involve concerns such as the voluntary participation of those involved, the entitlement of these people to participate, and their entitlement to whatever gains or losses the procedure's results might entail. This is certainly in line with how professional golf works. As I have already noted, Rawls requires that fair bets have zero expectation of gain, which is a constraint that entitlement theorists like Nelson reject. Nelson maintains that if a person freely agrees to a bet that does not have zero expectation of gain, the gamble is justified on pure procedural grounds. Nevertheless, a bet involving sure loss—say a loss of fifty dollars on heads and a loss of sixty dollars on tails—hardly appears fair, even if a person freely agrees to it. An entitlement theorist might argue that such a bet is not justified on the grounds that no rational person could agree to it, but then this relies more on an account of rationality and less on one of entitlements. This suggests that entitlements do not exhaust the range of pure procedural justifications.

¹⁰The analysis of fair procedures in David M. Estlund, *Democratic Authority* (Princeton: Princeton University Press, 2008), pp. 65–97, for instance, seems to presume that pure procedure approaches must remain silent about outcomes. *ibid.*, pp. 79–80 contends that social disagreements limits the background circumstances to a condition of 'full anonymity', which states that a procedure must be 'blind to all features of the individuals in question' and that 'its results would not be different if any features of the relevant people were changed'. On this account, the only procedures that satisfy this are 'a random choice from alternative decisions . . . [.] a procedure linking the

the case of Jones and the bookie, background circumstances restrict Jones' permissible outcomes to a gain of five dollars when (the coin lands heads up), a loss of five dollars (the coin lands tails up), or no gain/loss (Jones declines the bet). Losing ten dollars from a single coin flip, for instance, is not permissible under these conditions. In this sense, background circumstances place independent and prior constraints on permissible outcomes. They are best understood as second-order criteria with respect to outcomes, whereas perfect and imperfect justifications appeal to first-order criteria that judge outcomes directly.¹¹ These second-order criteria need not identify a unique outcome, though in certain situations they might.¹² It would be a mistake, however, to view the range of permissible outcomes as all equally acceptable. The bookie cannot simply take Jones' five dollars, claiming that, after all, this is a permissible possibility. The procedure must actually be carried out—the coin must be flipped—to determine which permissible outcome is the acceptable one.¹³

This last point stresses that an important characteristic of pure procedural justifications is that the procedure must actually be carried out. If the sole value of a criminal trial lies in its ability to correctly ascertain the defendant's guilt, then there seems little reason to actually have a trial when everyone already knows that the police have admissible and irrefutable evidence of guilt. A criminal trial, however, arguably has some pure procedural value in not finding the defendant guilty without due process of law. If so, the trial must go on, giving the defendant the opportunity to be heard. Even if the outcome is a foregone conclusion, a pure

outcome to the weather, or any other fact or process external to the features of the relevant people; Estlund, *Democratic Authority*, pp. 79–80. Estlund's claim is that appealing to anything other than full anonymity involves process-independent, that is, substantive, standards. He is not entirely clear on why this must be the case, though the assumption seems to be that a pure procedural justification cannot appeal whatsoever to the types of outcomes a procedure should generate. This is suggested by his method for showing that social choice theory and 'deep deliberative democracy' are not purely procedural, *ibid.*, pp. 73–76, 87–93. If that is the case, this is an extremely demanding restriction. Even Estlund's condition of full anonymity fails to pass because even it refers to the types of outcomes that a procedure should return; namely, outcomes that would not vary if the features of the relevant people were changed.

¹¹Thomas Christiano, 'Estlund on Democratic Authority', *Journal of Political Philosophy* 17, 2 (June 2009): 228–240, pp. 229–233 makes a similar point when addressing David Estlund's argument (see n. 10).

¹²As I argue in section 2.3.1 (see n. 77) I believe that Rawls' derivation of his two principles of justice may be of the sort that does identify a unique outcome.

¹³This is not to presume that all procedures must ultimately select a uniquely acceptable outcome. Depending on how the procedure is defined, it may be possible for it to declare that several outcomes are equally acceptable. Ties may be permissible outcomes as well, as it were. For instance, a tie may result in a coin flip should the coin land perfectly balanced on its edge, though admittedly that is extremely unlikely. If background conditions require a unique outcome, then the procedure may need to include mechanisms for breaking ties, such as flipping the coin again.

procedural approach still requires actually carrying out the procedure.¹⁴

This rather lengthy analysis suggests that a pure procedural justification has five components: the specification of (1) what the procedure involves, (2) the contexts in which employing that procedure is appropriate, and (3) the relevant background circumstances necessary for a correct or fair procedure to apply in those contexts, along with (4) an account of how the given procedure is correct or fair in light of these circumstances and (5) the stipulation that that procedure must actually be carried out for any outcomes to also be correct or fair.¹⁵

For the remainder of this chapter, I consider various procedural justifications of democratic processes. Many, if not most, have both imperfect and pure aspects, but I keep these two dimensions distinct as much as possible in my discussion, since each provides significantly different reasons for the liberal state to employ democratic processes. I begin with two traditional imperfect procedural justifications of democratic processes and then follow a movement that slowly gives way to more pure procedural justifications.

¹⁴Saying that ‘the procedure must actually be carried out’ seems to implicitly involve an important counterfactual claim. Take the following example adapted from Estlund, *Democratic Authority*, p. 75: Suppose the bookie decides in advance to take Jones’ five dollars, regardless of what happens with the coin flip. As it turns out, however, Jones agrees to the bet and the coin lands tails, so Jones freely gives the bookie the money. Certainly they appear to have followed the procedure, but this is only coincidental—had the coin landed heads up, the very same distribution of money would have obtained, though it is then obviously unfair. This shows that the advertised procedure was not carried out and an altogether different one was followed, though in this particular instance Jones cannot distinguish this fact. Only considering the counterfactual outcomes of the coin flip reveals this. As Estlund says, ‘the outcome of a fair coin flip . . . is . . . fair only if the outcome is produced by the coin flip in the right way’. This shows that a pure procedural approach assumes some non-accidental connection between the outcomes and the procedure, though the precise character of this connection is not entirely clear. I admit to being somewhat coy about this connection. Estlund is less so: he explicitly speaks of a ‘causal relation’ between the procedure and the outcome, though he provides no specifics for how such a relation must operate. Regardless, this claim might be justified by a counterfactual theory of causality such that of David Lewis, ‘Causation’, *Journal of Philosophy* 70, 17 (October 1973): 556–567. I am less inclined to argue this because an extensive literature shows that examples like the one involving Jones and the bookie raise a wide variation of competing intuitions concerning causation. See, for instance, Christopher Hitchcock, ‘Of Humean Bondage’, *British Journal for the Philosophy of Science* 54, 1 (2003): 1–25 and Clark Glymour, David Danks, Bruce Glymour, Frederick Eberhardt, Joseph Ramsey, Richard Scheines, Peter Spirtes, Choh Man Teng and Jiji Zhang, ‘Actual Causation: A Stone Soup Essay’, *Synthese* (forthcoming).

¹⁵Rawls, *Theory of Justice*, pp. 201, 362 also mentions ‘quasi-pure’ procedural justifications. I address these in section 2.3.1 in the context of Rawls’ theory.

2.2 Epistemic Conceptions of Democracy

Following Jules Coleman and John Ferejohn, I call an account that presents an imperfect procedural justification of democratic processes an epistemic conception of democracy.¹⁶ Such an explanation involves three components. First, it specifies a standard, independent from the outcomes of democratic processes, for assessing state decisions and contends that this is indeed the appropriate standard for doing so. Second, it argues that, under the right background circumstances, democratic processes are more or less reliable means for making good decisions according to that standard. So even if there are different individual assessments concerning which available option is best according to the standard, the decisions of democratic processes provide the best evidence for which assessment is correct. Third, it explains why democratic processes are necessary, given the independent standard. For, on the one hand, if everyone is equally competent at applying the relevant standard, democratic processes are unnecessary; any randomly selected person could make state decisions.¹⁷ If, on the other hand, individual expertise varies, then following Socrates' argument in the *Republic*, an oligarchy of experts appears superior to democratic processes.¹⁸

An epistemic conception of democracy should address these concerns. If successful, such an account should demonstrate why each of the state's constitutive

¹⁶Jules Coleman and John Ferejohn, 'Democracy and Social Choice', *Ethics* 97, 1 (October 1986): 6–25. Also see Joshua Cohen, 'An Epistemic Conception of Democracy', *Ethics* 97, 1 (October 1986): 26–38; David M. Estlund, 'Democracy Without Preference', *Philosophical Review* 99, 3 (July 1990): 397–423, 'Making Truth Safe for Democracy', in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 72–100, 'Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 173–204, 'The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth', *Ethics* 108, 2 (January 1998): 252–275, *Democratic Authority*; Christian List and Robert E. Goodin, 'Epistemic Democracy: Generalizing the Condorcet Jury Theorem', *Journal of Political Philosophy* 9, 3 (September 2001): 277–306; and Elizabeth S. Anderson, 'The Epistemology of Democracy', *Episteme* 3, 1–2 (June 2006): 8–22.

¹⁷This is a point used against epistemic accounts made by William H. Riker, *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice* (San Francisco: W. H. Freeman, 1982), 291 n2.

¹⁸*Republic*, book IV [435c–444e] in Plato, *Complete Works*, ed. by John M. Cooper and D. S. Hutchinson (Indianapolis, IN: Hackett, 1997), pp. 1067–1076. Along more democratic lines, Mill's *Considerations on Representative Government* (1861), chapter VIII in volume 19 of John Stuart Mill, *Collected Works of John Stuart Mill*, ed. by J. M. Robson (Toronto: University of Toronto Press, 1963–1991), pp. 473–479, proposes general suffrage but argues for a system of 'plural voting' whereby that the votes of those with higher education or intelligence should count for more. For a contemporary proposal along these lines, see Bryan Caplan, *The Myth of the Rational Voter: Why Democracies Choose Bad Politics* (Princeton: Princeton University Press, 2007).

members have a compelling reason to accept the decisions of democratic processes, even if some disagree with its outcomes in particular instances. The assumption is that these decisions should provide sufficient evidence for convincing a reasonable person as to which element is the right one according to the recognized standard. The analogy here is with the criminal justice system: some might disagree with the outcome of a criminal trial, but the process provides some assurances that it gave the correct result while also offering a chance to have the decision reconsidered through an appeal. Applied to democratic processes, dissenters should have both reasonable assurance that these outcomes are in accord with the recognized standard and opportunities to reconsider the issue in the future. If a reasonable person must accept that the decisions of a constitutional democracy are most likely correct, then the liberal test of legitimacy seems to have been met.

This last move to legitimacy is where the problem arises. Reasonable debates over the relevant standards with which to assess state decisions poses a serious challenge to epistemic views of democracy. In this section, I present two contrasting epistemic approaches that are illustrative of this point. Both are ultimately unsatisfying because they have problems reconciling their respective epistemic standards with the fact of reasonable pluralism. Even so, these two accounts have critical elements that more pure procedural justifications of democracy, which I address later, draw upon.

2.2.1 Bentham's Utilitarian Justification of Democracy

Jeremy Bentham provides a particularly straightforward example of an epistemic conception of democracy.¹⁹ Characteristic of a utilitarian, he maintains that 'the right and proper end of government in every political community, is the greatest happiness of all the individuals of which it is composed, say, in other words, the greatest happiness of the greatest number'.²⁰ As he understands it, the common good of the constitutive members of the state is happiness, which is the mental state associated with pleasure and the absence of pain. This notion of the common good is consistent with societal pluralism insofar as it recognizes that different individuals may find happiness in different things. Even so, Bentham maintains that happiness is the common denominator and that this requires the state to promote general utility, which is the aggregate sum total of happiness possessed by the constitutive members of the state. In this way, Bentham endorses a form of the

¹⁹My analysis of Bentham in this section follows themes presented in Ross Harrison, *Bentham* (London: Routledge and Kegan Paul, 1983).

²⁰*Constitutional Code* (1830, 1843), book I, introduction, section 2 in volume 9 of Jeremy Bentham, *The Works of Jeremy Bentham*, ed. by John Bowring (London: Simpkin, Marshall, and Company, 1838–1843), p. 5.

principle of equal consideration: everyone is equally entitled to the state's regard insofar as no one's pleasure or displeasure counts for more than anyone else's in the calculation of general utility.²¹

Calculating general utility seems to provide a perfect procedural approach for assessing state decisions that has nothing to do with democratic processes. It seems that a benevolent dictator could calculate for herself which laws and decisions maximize general utility and rule the state accordingly. According to Bentham, however, making such a calculation is extremely demanding. Doing so requires (1) determining the extent to which each option produces happiness in each of the state's constitutive members and (2) applying a metric for interpersonal comparisons of individual utilities in order to aggregate this information into a single measure of general utility. The calculation also requires substantial information concerning probability assessments for the various consequences that might result and determine how these balance out over a long period of time.

In light of these informational demands, Bentham believes that it is not feasible for the state to directly apply the principle of utility. There are simply no experts who could perform the necessary calculations. The subjective nature of this conception of happiness leads Bentham to adopt a presumption of personal autonomy maintaining that, absent strong evidence to the contrary, a person is presumed to be (1) the best judge of what makes her happy and (2) the least inclined to overlook her own happiness.²² On this account, no one can reliably know the interests of a person without first consulting her and no one else can more reliably seek to promote these interest than that person herself. This reasoning sparks suspicion in Bentham of professed experts who would impose their views on the public in the name of general utility, for '[no] man [is] now so far elevated above his fellows, as that he should be indulged in the dangerous license of cheating them for their own good'.²³ Together, the principle of equal consideration and the

²¹*Plan of Parliamentary Reform* (1817), introduction, section 7 in volume 3 of Bentham, *Works*, p. 459. Mill summarizes this principle as 'Bentham's dictum' that 'everybody to count for one, nobody for more than one', *Utilitarianism* (1861), chapter 5 in volume 10 of Mill, *Collected Works*, p. 257. I call this the principle of equal consideration based on Robert A. Dahl, *Democracy and Its Critics* (New Haven, CT: Yale University Press, 1989), p. 85, though Dahl uses it in a different context. See section 2.4.

²²*Institute of Political Economy* (1801–1804), The Art, part I, section 2 in volume 3 of Jeremy Bentham, *Jeremy Bentham's Economic Writings*, ed. by Werner Stark (London: George Allen and Unwin, 1952–1954), p. 333 and *An Introduction to the Principles of Morals and Legislation* (1789), chapter XVIII, section 3:44 in volume 1 of *Works*, p. 124. Mill's *Considerations on Representative Government* (1861), chapter III in volume 19 of Mill, *Collected Works*, pp. 404–406, endorses part (2) of this presumption while giving rather tepid support to (1). This probably explains why Mill advocates his system of plural voting (see n. 18). Once again, I follow Dahl, *Democracy and Its Critics*, p. 100 in calling this the presumption of personal autonomy.

²³*A Fragment on Government* (1776, 1822), chapter I, section 37 in volume 1 of Bentham, *Works*,

presumption of personal autonomy therefore lead Bentham to deny that experts should make decisions for the state.

Instead of experts, Bentham posits that democratic institutions, and majority rule in particular, are more reliable means for the state to achieve results that maximize general utility: 'the surest visible sign and immediate evidence of general utility is general consent', where the consent of a majority is sufficient for general consent.²⁴ The idea is that a person reliably assesses the available options and orders them according to how conducive they are to her own happiness. However, in a pluralistic society, people value different things because they find happiness in different things. Majority rule is then supposed to reveal any overlap between these competing sources of happiness insofar as it selects the outcome that makes the greater number of people happier. Bentham argues that this outcome usually coincides with the option that maximizes general utility, most likely because he believes that allowing people to vote provides a more complete reading of how much happiness each person contributes to the general utility under each option. Circumventing popular participation, on the other hand, would require using less reliable estimates instead.

The value of democratic institutions in Bentham's account is entirely instrumental.²⁵ Should empirical results demonstrate that experts are more reliable at making decisions that promote general utility, this approach would support employing them instead. In fact, he does concede that there are two circumstances that reveal imperfections in democratic processes for maximizing general utility, but he does not believe that they justify placing experts in charge.

In the first case, Bentham recognizes that people can be mistaken about their interests, as when they overly discount long-term interests in favor of short-term gains. It is therefore worrisome that majority rule may bankrupt the state by favoring excessive spending and little saving.²⁶ This is not devastating for democratic institutions, according to Bentham, because people are supposed to be motivated to constantly promote their own happiness, which leads the majority, in the long run, to naturally self-correct mistakes in maximizing general utility.²⁷ Experts, meanwhile, have their own discounting problem, as they may, intentionally or not,

p. 269.

²⁴Bentham MSS at University College London (box 127, sheet 3).

²⁵Instrumental accounts of democracy need not be limited to utilitarian ones. My account of Rousseau's theory of democracy in the next section emphasizes its instrumental aspects. More contemporary instrumental approaches include Richard J. Arneson, 'Defending the Purely Instrumental Account of Democratic Legitimacy', *Journal of Political Philosophy* 11, 1 (March 2003): 122–132 and Steven Wall, 'Democracy and Equality', *Philosophical Quarterly* 57, 228 (July 2007): 416–438.

²⁶Bentham proposes an annuities scheme that is supposed to address this problem in *Circulating Annuities* (1800) in volume 2 of Bentham, *Economic Writings*.

²⁷*Principles of the Civil Code* (1843), part I, chapter xvii in volume 1 of *Works*, p. 324.

favor their own interests over those of others—and in this case, their motivation to pursue their own happiness compounds the problem.

In the second case, Bentham notes that the majority could favor a policy giving them slight gains that come at the expense of extreme losses for the minority.²⁸ If each person voted according to what promotes her own happiness, this policy might have majority approval—even though it may fail to maximize general utility. Bentham's framework handles this problem insofar as his utilitarian approach to constitutional codes does require the state to adopt liberal institutions that might prevent such options from being on the ballot.²⁹ Regardless, putting experts in charge does not seem to avoid this problem, they'd be free to give themselves gains at significant expense to everyone else, which seems even more likely to violate the principle of utility.

In these respects, Bentham's epistemic conception of democracy relies on certain contingencies—ones he believes are firmly rooted in the human condition—that favor democratic processes over experts for reliably promoting general utility. Democratic processes might not be perfect, but they currently provide the best evidence for what maximizing general utility requires.

Even so, few political theorists adopt Bentham's approach. Its problems as a moral theory are well known: the assumption that certain subjective mental states suffice to ground all morality is far too limiting and the focus on aggregate happiness conflates the distinct interests of multiple individuals into that of a single, social entity, to name just two.³⁰ These issues illuminate why Bentham's approach does not satisfactorily legitimize the outcomes of democratic processes. The principle of utility is part and parcel of an extremely comprehensive doctrine—as I noted in section 1.2.1, few doctrines are as inclusive, general, and systematic as classical utilitarianism. So even if a reasonable person should agree that democratic decisions promote general utility, she may reasonably reject to assessing state decisions by this standard.

The upshot of all this is that it is not enough to wrap a comprehensive doctrine

²⁸Bentham MSS at University College London (box 14, sheet 386), published as part of *Article on Utilitarianism: Long Version*, sections 54–57 in Jeremy Bentham, *Deontology together with A Table of the Springs of Action and the Article on Utilitarianism*, ed. by Amnon Goldworth (Oxford: Oxford University Press, 1983), pp. 309–310.

²⁹In *Anarchical Fallacies* (1843), article II in volume 2 of *Works*, p. 501, Bentham rather famously declares that 'natural and imprescriptible rights, rhetorical nonsense,—nonsense upon stilts'. Despite believing that there is no law of nature with which to justify rights, Bentham does argue that personal and political liberties, along with security and a certain level of material equality, are instrumental in promoting general utility. See *Principles of the Civil Code* (1843), part I in volume 1 of *ibid.*, pp. 301–326.

³⁰These two criticisms of utilitarianism are respectively from Robert Nozick, *Anarchy, State, and Utopia* (New York: Basic Books, 1974), pp. 42–45 and Rawls, *Theory of Justice*, pp. 22–33.

in the garb of liberal democracy in order to satisfy the liberal test of legitimacy. An epistemic conception of democracy must provide an independent standard whose appropriateness for assessing state decisions cannot be reasonably denied. This poses a serious challenge for this type of approach, as the fact of reasonable pluralism raises doubts about the existence of any generally accepted standard whose satisfaction would ensure that democratic processes result in legitimately binding decisions.

2.2.2 Rousseau and the Condorcet Jury Theorem

Bentham's defense of democratic processes is certainly problematic, but it is appealing insofar as it shows how these processes might reliably ensure that state decisions satisfy some independent political standard. Jean-Jacques Rousseau's approach retains this appealing idea while providing an alternative epistemic justification for democratic processes that does not seem to require acceptance of a controversial comprehensive doctrine. Despite that, clarifying his account reveals that it too has problems similar to those facing Bentham's.³¹

Like Bentham, Rousseau argues that state decisions should be assessed by whether or not they promote the common good shared by all of the state's constitutive members. As Rousseau understands it, though, this common good is neither happiness nor maximizing general utility. Rousseau agrees with Bentham that each constitutive member of the state has a private will that is often motivated by its own personal interests. But while a person's personal interests may provide sufficient grounds for justifying that person's particular actions, these reasons remain personal because they do not necessarily motivate other people to act. As a result, Rousseau denies that personal reasons are appropriate reasons for justifying collective action organized by the state and he dismisses general utility as 'only a sum of private wills'.³² Instead of maximizing the overall extent to which the private wills of the state's constitutive members are satisfied, Rousseau argues that state action is supposed to be guided by a 'general will'. He claims that this will is motivated by fundamental interests that the state's members hold in common, which are, at bottom, assurances of 'the goods, life, and freedom of each member

³¹The account of Rousseau I present here is influenced by Joshua Cohen, 'Reflections on Rousseau: Autonomy and Democracy', *Philosophy and Public Affairs* 15, 3 (Summer 1986): 275–297; Bernard Grofman and Scott L. Feld, 'Rousseau's General Will: A Condorcetian Perspective', *American Political Science Review* 82, 2 (June 1988): 567–576; and John Rawls, *Lectures on the History of Political Philosophy*, ed. by Samuel Freeman (Cambridge, MA: Belknap Press of Harvard University Press, 2007), pp. 188–248.

³²*On the Social Contract*, book II, chapter iii in Jean-Jacques Rousseau, *On the Social Contract, with Geneva Manuscript and Political Economy*, ed. by Roger D. Masters, trans. by Judith R. Masters (New York: St. Martin's Press, 1978), p. 61.

by the protection of all'.³³

Rousseau's account suggests that the common good involves some subset of overriding interests and capacities that all people share, which does not rest on a shared desire to experience certain mental states.³⁴ Furthermore, Rousseau believes that the state cannot sacrifice these common interests to further the personal interests of some individuals. Therefore, when the state rules in accord with a constitution consistent with the general will, it protects and promotes the fundamental interests of each and every one of its constitutive members. This is why Rousseau claims that these people remain free—that is, they each obey their own wills by securing their own fundamental interests—even when being obedient to the state.³⁵ As a result, decisions made in accord with the general will may satisfy the liberal test of legitimacy, for a person cannot reasonably object to obeying her own will.

Beyond these rather broad strokes, Rousseau's account of the common good lacks the relative precision of Bentham's formulation, owing to the former apparently lacking the inclusiveness, generality, and systematicity of the latter. For instance, Rousseau does not explain whether it is possible for fundamental interests to conflict, and, if so, how they ought to be traded off against each other. In addition to this lack of systematicity, Rousseau's conception of the common good does not specify many determinate positions on many of the controversial issues that the state must address. Apart from the need of the state to protect a person's goods, life, and liberty, Rousseau apparently leaves it to individual societies to fill in the particulars on their own, though he does propose how particularities of different societies, such as geographical location and the characteristics of its inhabitants, should influence their respective understandings of the common good.³⁶

This last point suggests that Rousseau appears to assume that the constitutive members of the state share a conception of the common good, though some of its

³³*Discourse on Political Economy* in Jean-Jacques Rousseau, *The First and Second Discourses*, ed. by Roger D. Masters, trans. by Roger D. Masters and Judith R. Masters (New York: St. Martin's Press, 1964), p. 214. Also see *On the Social Contract*, book II, chapters i, xi in *On the Social Contract*, pp. 59, 75.

³⁴This reading is lent further support from Rousseau's comparison between animals as slaves to instinct and humans as free to reject it in the *Discourse on the Origin and Foundations of Inequality* in *First and Second Discourses*, p. 113. This understanding of Rousseau is also influenced by the taxonomy of conceptions of the common good presented in Alex John London, 'Threats to the Common Good: Biochemical Weapons and Human Subjects Research', *Hastings Center Report* 33, 5 (September 2003): 17–25. According to this taxonomy, Bentham adopts a 'corporate' conception of the common good whereas Rousseau has a 'generic interests' conception.

³⁵*On the Social Contract*, book II, chapter iv in Rousseau, *On the Social Contract*, p. 63.

³⁶*On the Social Contract*, book II, chapter xi in *ibid.*, pp. 75–76.

specific contents may vary from state to state. Societal pluralism results on this account because the members of a given state may disagree on how these contents apply to particular decisions. On this account, however, such disagreements have their sources in weakness of mind or weakness of will. Individuals are prone to weakness of mind because a person often lacks either the information about proposed social policies or the time to process and evaluate it, thus leading to errors in judgment concerning the common good. Weakness of will is also possible due to the fact that a person's personal interests may diverge from those of the common good, tempting her to pursue social arrangements that benefit the former.³⁷

Since personal interests and misinformation may cloud individual judgments concerning the common good, Rousseau denies that experts ought to make state decisions. Given the reins of state power, anyone would be tempted to rule in accord with her personal interests and not with the common good. Those subject to the rule of an oligarchy, even one of putative experts, should reasonably worry whether a given decision was made according to the common good. Instead, Rousseau argues that democratic processes, again with the emphasis on majority rule, can reliably track the dictates of the general will.³⁸ As I understand his account, Rousseau believes that these institutions ensure, under the right conditions, that correct assessments of the common good wash out mistaken ones. While this process may not be perfect, it provides the best evidence for what the general will requires. Consequently, those who oppose a decision made by democratic institutions may then be accused of either making an error in judgment or demanding that the state sacrifice the common good in order to satisfy their own personal interests.³⁹ This suggests why such people have little basis for reasonably opposing the majority's decision, showing how the liberal test of legitimacy might be satisfied.

Rousseau's advocacy of democratic processes may seem initially puzzling. He appeals to very same processes that Bentham does for decisions making by the state, and yet Rousseau wants to claim that they provide evidence for something that is very different from what Bentham supposes. This raises an important question: How can the very same procedures reliably track two mutually incompatible notions? For, on the one hand, Bentham maintains that democratic processes promote general utility; Rousseau, on the other hand, argues that they reveal the

³⁷ *On the Social Contract*, book I, chapter vii in *ibid.*, p. 55.

³⁸ *On the Social Contract*, book IV, chapter ii in *ibid.*, pp. 109–111. This is not to suggest that Rousseau's account of democracy is solely epistemic. There are surely other reasons, according to Rousseau, for instituting democratic processes, as suggested in Melissa Schwartzberg, 'Voting the General Will: Rousseau on Decision Rules', *Political Theory* 36, 3 (June 2008): 403–423.

³⁹ *On the Social Contract*, book IV, chapter ii in Rousseau, *On the Social Contract*, pp. 110–111.

general will and that this is not general utility.⁴⁰ I believe that this difference is best explained by reference to the alternative sets of background circumstances for democratic processes that each thinker specifies.

According to Bentham, the common good is maximizing general utility, which is the aggregate satisfaction of personal interests. This leads Bentham to suggest background circumstances that require each individual to vote according to her own private interests, with majority rule thought to usually leading to outcomes that promote aggregate utility. On Rousseau's conception, the common good is not an aggregative concept at all; it is rather securing something that each of the constitutive members of the state are supposed to hold in common. While Rousseau does provide a much more extensive range of background circumstances than Bentham, which I discuss momentarily, Rousseau himself is not entirely clear on how these are supposed to ensure that democratic processes reliably promote the common good.

Some recent commentators suggest that the Jury Theorem—a mathematical result by Rousseau's contemporary, the Marquis de Condorcet—can credibly explain Rousseau's epistemic approach to democracy.⁴¹ While Rousseau probably did not know about the Jury Theorem when he wrote *On the Social Contract*, the idea behind it goes back at the very least to Aristotle's claim that a group may often make a better decision than any of its constituent members on her own.⁴² The intuition is that individuals may have imperfect information and

⁴⁰In 1792, Bentham was offered an honorary citizenship to the new French republic. In his acceptance letter, Bentham interestingly identifies the 'general will' with the 'general good', the latter of which for Bentham certainly means general utility. See chapter XI of Bentham's memoirs and correspondences in volume 10 of Bentham, *Works*, p. 282.

⁴¹Formal presentations of the Jury Theorem can be found in Marquis de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris: L'Imprimerie Royale, 1785), the relevant sections of which may be found in 'From *An Essay on the Application of Analysis to the Probability of Decisions Rendered by a Plurality of Votes*, 1785', in *Classics of Social Choice*, ed. and trans. by Iain S. McLean and Arnold B. Urken (Ann Arbor: University of Michigan Press, 1995): 91–111; Duncan Black, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958), pp. 159–180; and Bernard Grofman, Guillermo Owen and Scott L. Feld, 'Thirteen Theorems in Search of Truth', *Theory and Decision* 15, 3 (September 1983): 261–278. On the application of the Jury Theorem to Rousseau's theory, see Brian Barry, 'The Public Interest', *Proceedings of the Aristotelian Society* 38 (1964): 9–14; Cohen, 'An Epistemic Conception of Democracy'; Grofman and Feld, 'Rousseau's General Will: A Condorcetian Perspective'; and David M. Estlund, Jeremy Waldron, Bernard Grofman and Scott L. Feld, 'Democratic Theory and the Public Interest: Condorcet and Rousseau Revisited', *American Political Science Review* 83, 4 (December 1989): 1317–1340, Estlund, 'Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority'.

⁴²'Now any member of the assembly, taken separately, is certainly inferior to the wise man. But the state is made up of many individuals. And as a feast to which all the guests contribute is better than a banquet furnished by a single man, so a multitude is a better judge of many things than any

circumscribed perspectives, but when they come together as a group, they can pool their information and share their perspectives in order to make more accurate judgments.⁴³ In this way, voting need not be regarded simply as a process for aggregating personal preferences concerning a set of options, as it is for Bentham. Instead, voting can be part of a cognitive enterprise for assessing options with respect to how they further the common interests of those involved and arriving at a considered judgment concerning which option does that best.⁴⁴

The Jury Theorem demonstrates how majority rule can have this effect. Roughly speaking, suppose that a competent person is more likely than not to give the correct answer to a particular yes or no question. As a group of competent people gets larger, the Jury Theorem reveals how majority rule quickly converges on the right answer to that question. For this to happen, the Jury Theorem requires five important stipulations:

A Well-Formed Question: Each voter renders judgment concerning the same question, which has a correct answer that is independent of the vote's outcome.⁴⁵

Widespread Participation: There must be a significant number of people who answer the question.

Voter Competence: Each voter has greater than a fifty percent chance of giving the correct answer.

Voter Independence: Each voter arrives at her own answer independently of the other voters insofar as each does not unreflectively defer to the judgment of another.

Binary Choice: There are only two possible answers (yes or no, candidate *A* or candidate *B*, and so forth).

individual', *Politics*, book III, chapter 15 [1286^a28–31] (trans. by Benjamin Jowett) in volume 2 of Aristotle, *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. by Jonathan Barnes (Princeton: Princeton University Press, 1984), p. 2041.

⁴³Empirical studies show that in certain situations groups are often more reliable than any one of their constitutive members, such as when guessing the number of jellybeans in a jar or determining the odds in sports betting. For a broad survey of this 'wisdom of crowds', see James Surowiecki, *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business* (New York: Doubleday, 2004). For a collection of article on how to model information pooling, see Bernard Grofman and Guillermo Owen, eds., *Information Pooling and Group Decision Making* (Westport, CT: JAI Press, 1986).

⁴⁴Cohen, 'An Epistemic Conception of Democracy', p. 34.

⁴⁵In many presentations of the Jury Theorem, this condition is left largely implicit. For reasons that will become apparent shortly, it is crucial that this condition is not ignored.

Recent results show how the last three of these can be weakened in certain ways, though the results are not nearly as striking.⁴⁶

The background circumstances for democratic processes that Rousseau specifies seem to have something like the Jury Theorem in mind. The need for a well-formed question and binary choice appears in Rousseau's demand that voters restrict themselves to considering whether a given proposal does or does not promote their shared conception of the common good.⁴⁷ His requirement that voters are 'adequately informed' raises voter competency.⁴⁸ Rousseau also repeatedly emphasizes the need for a social bond between people that is not strained by extreme differences in wealth or polarizing societal fractions.⁴⁹ This further raises voter competency by encouraging people to vote in accord with what they actually believe is the common good and not with their private interests. It also ensures voter independence, for individuals then make up their own minds on the issue and not simply follow their political party or faction. In addition, Rousseau advocates widespread—even compulsory—participation in the process.⁵⁰ Finally, Rousseau recognizes that the form of voting may vary, depending on the issue at hand: bare majorities are sufficient for 'business matters' while the passage of important laws requires something closer to unanimity.⁵¹ This seems to suggest that the more important an issue, the more confident the state needs to be that it has made the right decisions, and the Jury Theorem explains how a bigger majority provides this assurance.

The Jury Theorem certainly provides a viable explanation for how democratic processes may reliably promote some non-aggregative conception of the common good. Nevertheless, the requirement that the state's constituent members share such a conception seems in tension with the fact of reasonable pluralism. Even though Rousseau is not clear on the determinate conception of the common

⁴⁶First, voter competency is not necessary as long as average voter competency remains above a fifty percent chance of answering correctly, see Theorems V–XI in Grofman, Owen and Feld, 'Thirteen Theorems in Search of Truth', pp. 268–272 and Theorems 2–4 in Philip J. Borland, 'Majority Systems and the Condorcet Jury Theorem', *Statistician* 38, 3 (1989): 181–189, pp. 183–184. Second, voter independence may be weakened to allow for some mild interdependencies between the choices of voters, see Theorem XII in Grofman, Owen and Feld, 'Thirteen Theorems in Search of Truth', pp. 273–274 and Theorems 5 and 6 in Borland, 'Majority Systems and the Condorcet Jury Theorem', pp. 185–186. Third, binary choice may be dropped, permitting decisions concerning more than two options, and there are Jury Theorems for these cases, as well, though they may require more widespread participation to compensate, see Propositions 1 and 2 in List and Goodin, 'Epistemic Democracy: Generalizing the Condorcet Jury Theorem', pp. 295–297.

⁴⁷*On the Social Contract*, book IV, chapter ii in Rousseau, *On the Social Contract*, pp. 110–111.

⁴⁸*On the Social Contract*, book II, chapter iii in *ibid.*, p. 61.

⁴⁹*On the Social Contract*, book II, chapters iii, xi and book IV, chapter i in *ibid.*, pp. 61, 75, 108.

⁵⁰*On the Social Contract*, book III, chapter xv in *ibid.*, pp. 101–104.

⁵¹*On the Social Contract*, book IV, chapter ii in *ibid.*, p. 111.

good, he apparently envisages fairly homogenous societies where there is little or no reasonable disagreement concerning this conception, whatever it ends up being.⁵² Regardless, there must be agreement on the conception of the common good and a general recognition of this by voters in order for the Jury Theorem to work. Otherwise, the question ‘does this policy promote the common good?’ is not necessarily well-formed, and could mean different things to different people. For instance, a Benthamite would see this as asking about general utility while a libertarian would interpret it as asking about potential infringements on individual liberty. Stipulating a conception of the common good involved creates a well-formed question, but this only raises the concern that, given the fact of reasonable pluralism, people may reasonably reject this particular formulation of the common good as appropriate for justifying state action.

In this respect, Rousseau’s view faces a similar objection as Bentham’s account: the fact of reasonable pluralism denies that the state’s constitutive members must agree on an independent standard for assessing state decisions. In both instances, reasonable pluralism casts doubt on the plausibility of these traditional epistemic conceptions of democracy. This is regardless of whether the standard they appeal to is aggregative or not, and regardless of whether they rely on the Jury Theorem or something else when justifying how democratic processes reliably produce outcomes adhering to that standard. In order to succeed, an epistemic conception of democracy must provide an independent standard that cannot be reasonably denied by dissenting minorities and others. This problem leads many liberals to reject strongly epistemic approaches and move towards more purely procedural ones instead.

2.3 Contractualist Defenses of Democracy

Concerns with epistemic justifications of democracy raise the issue of whether a system of democratic institutions can reliably lead to outcomes promoting a process-independent standard such as the common good or justice when there is no widespread social agreement on what such a standard ought to be. In response, many contemporary liberals increasingly appeal to what Jürgen Habermas characterizes as a ‘postmetaphysical authority independent of comprehensive doctrines.’⁵³ As he puts it:

In pluralistic societies in which comprehensive worldviews and collect-

⁵²This seems a reasonable conclusion to draw from *On the Social Contract*, book II, chapter i in *ibid.*, p. 59.

⁵³Jürgen Habermas, *The Inclusion of the Other*, ed. by Ciaran P. Cronin and Pablo de Greiff, trans. by Ciaran P. Cronin (Cambridge, MA: MIT Press, 1998), p. 95.

ively binding ethics have disintegrated, societies in which the surviving posttraditional morality of conscience no longer supplies a substitute for the natural law that was once grounded in religion or metaphysics[,] the democratic procedure for the production of law evidently forms the only postmetaphysical source of legitimacy.⁵⁴

The tactic here is to emphasize that the legitimacy of a constitutional essential, a law, or a more particular decision depends on its provenance, on the features of the process that lead to it, and not on process-independent standards to which the decision should conform. Instead of arguing that democratic institutions reliably *discover* and promote the determinate conception of the common good or justice for society, the claim is that these institutions are integral in *fashioning* and promoting such a conception.

One step in this direction is to adopt a contractualist approach that models the liberal test of legitimacy under idealized circumstances.⁵⁵ That is, contractualism maintains that legitimacy derives from decisions made in an idealized collective choice situation by idealized persons acting as representatives for the state's constitutive members. Legitimate exercises of power are those that these people in that situation would agree to authorize. Call this the contractualist test of legitimacy. According to this test, legitimacy does not necessarily depend on

⁵⁴Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. by William Rehg (Cambridge, MA: MIT Press, 1996), p. 448. Compare to John Rawls, *Justice as Fairness: A Restatement*, ed. by Erin Kelly (Cambridge, MA: Harvard University Press, 2001), p. 15:

Given the assumption of reasonable pluralism, citizens cannot agree on any moral authority, say a sacred text or religious institution or tradition. Nor can they agree about a moral order of values or the dictates of what some view as natural law. So what better alternative is there than an agreement between citizens themselves reached under conditions that are fair to all?

⁵⁵The views that I have in mind here are *Theory of Justice, Political Liberalism*, (New York: Columbia University Press, 1993); William Nelson, *On Justifying Democracy* (London: Routledge and Kegan Paul, 1980); Joshua Cohen, 'Deliberation and Democratic Legitimacy', in *The Good Polity: Normative Analysis of the State*, ed. by Alan Hamlin and Philip Pettit (Oxford: Basil Blackwell, 1989): 17–34, 'Procedure and Substance in Deliberative Democracy', in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. by Seyla Benhabib (Princeton: Princeton University Press, 1996): 95–119; Jürgen Habermas, 'Discourse Ethics: Notes on a Program of Philosophical Justification', in *Moral Consciousness and Communicative Action*, trans. by Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: MIT Press, 1990): 43–115, *Between Facts and Norms, Inclusion of the Other*; and Brian Barry, *A Treatise on Social Justice*, vol. 2: *Justice as Impartiality* (Oxford: Oxford University Press, 1995). Thomas M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Belknap Press of Harvard University Press, 1998) is also extremely influential, though Scanlon presents a contractualist theory for general morality and not political theory.

what the state's real members actually agree to authorize; hypothetical agreements can be legitimately binding on actual, nonidealized people.⁵⁶

The circumstances of such a hypothetical agreement usually involve deliberations under what are supposed to be the most favorable conditions for a group of people to make a fair decision. Fleshing out the details of a contractualist approach requires specifying the characteristics of the deliberators and the circumstances of their deliberation. The first of these include the rationality, knowledge, and motivation of the parties seeking an agreement.⁵⁷ It also concerns the extent to which the deliberators admit of variations or are uniform in their characteristics.⁵⁸ The circumstances of their deliberation include considerations such as the deliberation's agenda, what occurs during its duration, and how the final agreement is determined.⁵⁹

⁵⁶Recall that in section 1.3 (see n. 76 in Chapter 1) I noted that this was one direction in which the liberal test of legitimacy might point.

⁵⁷Rationality in this context may be understood as referring, firstly, to the capability of a person to follow complex inferences and understand scientific theories. Secondly, it may refer to the effectiveness of a person in advancing her conception of the good or following her life plan. In this sense, a person is rational to the extent that she is capable of making her goals consistent with each other and taking up the effective means to these goals, which involves selecting those means that satisfy the most of her goals and that have a greater likelihood of success. Rawls, *Theory of Justice*, p. 415 calls these the 'counting principles.' This naturally leads to the question of how much a deliberator is supposed to know. She might have perfect knowledge about herself, her conception of the good life, her society, the consequences of adopting alternative courses of action, the psychology of persons, and the way the world works socioeconomically and scientifically. Or she could have average knowledge, with its usual imperfections, concerning these things. Finally, a contractualist theory must specify what is motivating the deliberators. One aspect of this specifies the interests they each want the agreement to promote. For instance, they might only be concerned to promote their respective personal interests and well-being. Another possibility is that they each seek to advance whatever it is that their respective conception of the good life involves, which may or may not be responsive to the interests and welfare of others. Or they might favor an agreement that equally promotes the interests and welfare of everyone involved. Another aspect of her motivation concerns the extent to which the parties are willing to abide by the terms of the agreement. They may each plan to violate them whenever given chance, they may adhere to them no matter what, or they may only follow these terms provided that everyone else does so.

⁵⁸If the deliberators are alike in all relevant respects, unanimous agreement is a trivial matter: once one person settles on an outcome, they all agree to it. On the other hand, if reasonable people vary in significant ways, unanimous agreement may not be possible, and so other methods for finalizing an agreement, by majority rule or something else, may be necessary.

⁵⁹The agenda of the deliberation concerns what exactly the agreement is about. That is, deliberators may be trying to reach an agreement on the state's constitutional essentials, or they may try to do so for each and every issue—from the basic structure of society to particular laws to their various applications—that affects the state. In either case, the agenda may also concern the principles that the state may appeal to for addressing social controversies, or it may concern the procedures that the state may employ for doing so. In all these various ways, the agenda of the deliberation determines how to assess state laws and decisions. The duration of the deliberation involves the issues concerning

While contractualist accounts may differ over the conditions for a legitimately binding hypothetical agreement, there do exist common strategies for using the contractualist test of legitimacy to support democratic institutions. In this section, I would like to consider two representative approaches. The first argues that a constitutional variant of the contractualist test of legitimacy requires democratic institutions (section 2.3.1). That is, deliberators would agree to include democratic institutions among the state's constitutional essentials and so legitimate exercises of state power must be done in accordance with the processes that these institutions require. This appears to be John Rawls' strategy. The second strategy maintains that democratic institutions more or less reliably produce decisions satisfying the contractualist test of legitimacy (section 2.3.2). This is the strategy that the 'deeply democratic' contractualist approaches of Habermas and Joshua Cohen adopt.⁶⁰ In the end, however, the contractualist test of legitimacy places legitimacy outside of what actual people have authorized the state to do, and a concern with this leads many liberals to advocate a liberal test of legitimacy that is not contractualist in this way.

2.3.1 Constitutional Contractualism

In his earlier work, John Rawls introduces a rather complicated contractualist approach. It involves a four-stage sequence of hypothetical agreements made between idealized representatives of the state's constitutive members, where the agreements at an earlier stage constrain the agreements of later stages.⁶¹ At the

what goes on during the deliberation. Deliberators may be able to introduce any alternatives for consideration or their decision may be constrained by a predetermined list of alternatives. Of course, there could instead be a more elaborate parliamentary procedure for introducing alternatives. A deliberator may be permitted to simply express her support or opposition to a given alternative under consideration, or she may be required to also present her reasons for her position to the other deliberators. If reasons are required, then it remains to specify the types of reasons a person may present. For instance, all reasons may count, or only those that acceptable to the others may be presented, or there may be predetermined restrictions on these reasons. A final concern is when is deliberation over and resolved. As I have mentioned already, the resolution of the deliberation may require unanimous agreement, or it may require something like a vote. If an unanimous agreement is required, then the concern is whether deliberation continue on, possible forever, as long as there is no unanimity. If a vote or other resolution process is permissible, then the concern becomes whether deliberation should be foregone and the process immediately employed once all the alternatives have been determined. If deliberation is still necessary, the account must then specify what must exactly happen before the resolution procedure may be employed.

⁶⁰I take the term deeply democratic from Frank I. Michelman, 'How Can the People Ever Make the Laws? A Critique of Deliberative Democracy', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 145–171, p. 152.

⁶¹Rawls, *Theory of Justice*, pp. 195–201.

first stage, in what Rawls calls the 'original position', the agenda asks deliberators to agree on principles governing the general arrangement of the state's political, constitutional, social, and economic institutions.⁶² Rawls argues that deliberators in this position would agree on his two principles of justice.⁶³ The next stage is a 'constitutional convention' where the agenda is now to agree on constitutional essentials that are not only consistent with the two principles of justice but also specify legislative processes for making laws that promote these two principles.⁶⁴ Rawls maintains that deliberators here would agree on an arrangement of liberal and democratic institutions.⁶⁵ After that, the legislative stage determines the laws that accord with these institutions and the two principles of justice.⁶⁶ Finally, there is the application stage, which concerns the application of the laws, institutions, and principles of justice to particular cases.⁶⁷

On this account, just exercises of state power are those that would be authorized at the application stage, though agreements at this stage are constrained by previous hypothetical agreements on principles of justice, the constitution, and the laws. At both this and the legislative stages, Rawls allows that unanimous agreement between even idealized deliberators under idealized conditions might not be possible. If disagreements over laws and applications remain, Rawls argues that these deliberators should employ democratic institutions such as majority rule in order to reach an agreement.⁶⁸ This is because at the earlier constitutional stage, deliberators would unanimously agree to use democratic processes to resolve such disagreements. Therefore, hypothetical agreements at the application stage, which set the standards for just exercises of state power, are made using democratic methods. In his later work, Rawls proposes a similar condition requiring that legitimate exercises of power by the state are those done in accordance with a constitution that deliberators would unanimously agree on at a hypothetical constitutional convention.⁶⁹ So Rawls's justification for democratic institutions rests on his argument that they pass the contractualist test of legitimacy.

While this standard for legitimacy involves an ideal democratic process, this alone does not yet require that real people must also use democratic processes

⁶²ibid., pp. 11–22. The agenda at this stage concerns what Rawls calls the 'basic structure of society'. For more on this, see John Rawls, 'The Basic Structure as Subject', *American Philosophical Quarterly* 14, 2 (April 1977): 159–165.

⁶³*Theory of Justice*, pp. 118–192.

⁶⁴ibid., pp. 196–198.

⁶⁵ibid., pp. 221–234.

⁶⁶ibid., pp. 198–199, 357–359.

⁶⁷ibid., pp. 199–200.

⁶⁸'Since even rational legislators would often reach different conclusions, there is a necessity for a vote under ideal conditions', ibid., p. 357.

⁶⁹Rawls calls this 'liberal principle of legitimacy', Rawls, *Political Liberalism*, p. 137.

in order to authorize exercises of state power. Imagine that a political savant programs a sophisticated computer simulation that can very reliably calculate which exercises of power deliberators at the application stage would authorize. Putting this computer program in charge of the state seems consistent with Rawls' contractualist standard of legitimacy. After all, when it comes to the original position and constitutional stages, Rawls claims to anticipate their outcomes, so it seems possible that similar experts might exist for the other stages. The short reply to this objection is that the hypothetical agreements at *each* step of the four-stage sequence apply to the actual state. According to Rawls, deliberators at the constitutional stage would agree to include democratic institutions among the constitutional essentials. This means that the constitutional variant of the contractualist test of legitimacy requires that actual exercises of power by the state must be done in accordance with democratic institutions, and this in turn seems to require that the state actually use democratic institutions.

This objection and its reply, however, raise two issues. First, Rawls claims to know a substantial amount of the content of the hypothetical agreements that would be made in the original position and at the constitutional stage. The concern is why anyone should take Rawls' word on this. Second, Rawls assumes that these hypothetical agreements have normative force for real people. This is because the determinants of legitimacy involve a hypothetical constitution that is independent from the agreements of actual people, and hence the worry that a computer simulation can make legitimate decisions regardless of what actual people may believe. To paraphrase Ronald Dworkin, it is mysterious how real people can be bound by agreements to which they were not actual parties; it is only something on which people *would* agree under certain special circumstances but not something on which they *have* agreed.⁷⁰ A hypothetical agreement hardly seems binding like an actual one, just as it would be wrong for a judge to forgo the due process of law and declare a defendant guilty—even if the judge rightfully recognizes that any judge or jury would find the evidence against the defendant conclusive.

Rawls responds by claiming that each stage of hypothetical agreement models considerations that real people accept and that these considerations allow him to both anticipate the content of the hypothetical agreements and show how these agreements are binding on real people. For instance, Rawls claims that most people, either already or upon philosophical reflection, agree that certain considerations should constrain the choice of principles governing the arrangement of the

⁷⁰Ronald Dworkin, 'The Original Position', *University of Chicago Law Review* 40, 3 (Spring 1973): 500–533, pp. 500–504. Also see A. John Simmons, 'Liberal Impartiality and Political Legitimacy', *Political Books* 34, 4 (October 1993): 213–223, pp. 220–221.

state's basic institutions.⁷¹ He maintains that: 'It seems reasonable and generally acceptable that no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles. It also seems widely agreed that it should be impossible to tailor principles to the circumstances of one's own case. We should insure further that particular inclinations and aspirations, and persons' conceptions of their good do not affect the principles adopted.'⁷² Rawls adds that the decision-making process 'must not permit some to have unfair bargaining advantages over others. Further, threats of force and coercion, deception and fraud, and so on must be ruled out.'⁷³

Considerations like these are modeled by the original position in two ways. First, all deliberators have an equal opportunity to make proposals, submit reasons justifying them, challenge the proposals of others, and otherwise participate in the process.⁷⁴ Second, the original position places all deliberators behind a 'veil of ignorance', which deprives them of knowledge concerning their respective races, genders, socioeconomic statuses, native talents and endowments, conceptions of the good life, and any other features that might distinguish them from each other. Furthermore, they are ignorant of the particulars concerning their society's resources, economic advantage, and level of culture.⁷⁵ These conditions put the deliberators into symmetrical positions of power, therefore preventing any of them from biasing the outcome in their own favor and nullifying the influence of any bargaining advantages.

Rawls maintains that given these and other features of the original position, deliberators there would agree must on his two principles of justice. If real people accept the considerations that the original positions models, the argument is that these people should regard agreements reached in the original position as binding, even though these agreements may remain entirely hypothetical.⁷⁶ The onus is on those who reject these agreements to demonstrate why the original position incorporates constraints that either are incomplete, fail to accurately reflect generally accepted intuitions concerning the selection of principles governing the arrangement of the state's basic institutions, or support a different agreement. Otherwise a dissenter is not reasoning consistently from considerations she already affirms. In this way, the original position attempts to model an argument for Rawls' two

⁷¹Rawls, *Theory of Justice*, p. 21.

⁷²ibid., p. 18.

⁷³Rawls, *Justice as Fairness: A Restatement*, p. 15.

⁷⁴*Theory of Justice*, p. 19.

⁷⁵ibid., pp. 136–142.

⁷⁶As ibid., p. 13 puts it: 'It will then be true that when whenever social institutions satisfy these principles those engaged in them can say to one another that they are cooperating on terms to which they would agree if they were free and equal persons whose relations with respect to one another were fair'.

principles of justice based on premises that are generally accepted.⁷⁷

Similar claims apply to the remaining stages in Rawls' sequence and therefore explain his justification of democratic institutions. According to Rawls, the agreement of deliberators at the constitutional stage on democratic institutions models a *quasi*-pure procedural justification for these institutions.⁷⁸ As I understand Rawls, a quasi-pure procedural justification is a hybrid between imperfect and pure procedural justifications. Its imperfect procedural (or epistemic) aspect argues that the procedure is reliable to a certain extent in ensuring better outcomes according to process-independent standards. However, these standards, for whatever reason, are indeterminate in that they only specify a set of outcomes that it judges as equally acceptable. The pure procedural aspect then argues that, in light of certain background circumstances, the procedure is the fair or correct way to actually choose an outcome from that set. In other words, the outcome chosen by the procedure is fair or correct because (1) it is reliable to a certain extent in picking among the acceptable outcomes according to the relevant independent standards and (2) the procedure itself is fair or correct in those circumstances.⁷⁹ A quasi-pure procedural justification therefore appeals to both outcome-dependent and outcome-independent aspects of the process.

Rawls provides such a justification for democratic institutions by arguing

⁷⁷This leads some to question whether Rawls' approach in *A Theory of Justice* is actually contractual. For instance, see Nozick, *Anarchy, State, and Utopia*, pp. 196n–197; Sidney S. Alexander, 'Social Evaluation through Notional Choice', *Quarterly Journal of Economics* 88, 4 (1974): 597–624; and Jean Hampton, 'Contracts and Choices: Does Rawls Have a Social Contract Theory?', *Journal of Philosophy* 77, 6 (June 1980): 315–338. Certainly the idea that the original position models an argument seems at odds with the claim that 'the original position is defined in such a way that it is a status quo in which any agreements reached are fair. . . . Thus justice as fairness is able to use the idea of pure procedural justice from the beginning', Rawls, *Theory of Justice*, p. 120. The problem is that purely procedural approaches require that the process be actually carried out. Perhaps Rawls believes this is a pure procedural approach where it just so happens that the background circumstances—the conditions modeled by the original position—entail a unique outcome. To 'actually carry out the procedure' in this context means to consider the argument for each background circumstance and how it is supposed to restrict the range of permissible outcomes. This is suggested by the statement that 'at any time we can enter the original position, so to speak, simply by following a certain procedure, namely, by arguing for principles of justice in accordance with these restrictions', *ibid.*, p. 19. The claim appears to be that accepting all the background circumstances necessarily entails that his two principles are the only permissible alternative.

⁷⁸*Ibid.*, pp. 201, 362.

⁷⁹Given my discussion of the role that background circumstances play in purely procedural approaches in section 2.1, arguably a purely procedural justification might be put in quasi-pure procedural terms by converting second-order constraints on outcomes into first-order ones. This could be done with fair gambling, where the background circumstances may be reinterpreted as setting process-independent standards delineating a set of permissible distributions of money. Even so, the gambling process must still be carried out to determine which of the possible distributions of money is the acceptable one.

that they (1) are themselves just and (2) reliable to a certain extent in producing just outcomes, where 'just' in both cases is determined by the two principles of justice. Starting with the first part, Rawls claims that the first principle of justice requires that the state's constitutive members have equal political liberty, which is embodied by a principle of equal participation claiming that all of the state's constitutive members must have an equal opportunity to participate in and influence the outcomes of political decisions.⁸⁰ This influence may either be direct through voting or indirect through one's duly elected representative. In either case, this opportunity must have its 'fair value', according to Rawls, in the sense that inequalities due to one's socioeconomic status do not determine one's political influence, either in voting or in running for elected office.⁸¹ Ensuring the fair value of equal participation is therefore a background circumstance for a just decision-making process. Putting a computer simulation or putative experts in charge of the state is incompatible with this condition and so Rawls concludes that only an arrangement of democratic institutions with supporting liberal institutions satisfies it. As a result, these institutions have just features that are irrespective of the outcomes they may produce.

Even so, the second, epistemic, part of Rawls' quasi-pure procedural justification argues that liberal and democratic institutions should be arranged so that they reliably produce just outcomes as delineated by the two principles of justice. The second principle, for instance, requires fair equality of opportunity, whereby those with similar abilities and ambitions should have similar prospects of success in attaining desirable positions and careers within society, and the difference principle, which claims that inequalities attached to these positions should be arranged so that they are to the greatest benefit of the least advantaged.⁸² Conditions like these provide process-independent standards of justice, though they may only denote a range of equally just laws and decisions.⁸³ Rawls believes that actual democratic institutions can be arranged so as to be somewhat reliable in selecting outcomes that fall within this range, though it is not entirely clear why this is the case. The claim is that democratic processes are a good way for 'combining information and enlarging the range of arguments' in order to ensure that outcomes are just.⁸⁴ Furthermore, certain democratic arrangements, such as propositional representation, along with constitutional guarantees of basic liberties, may help guide decisions towards just outcomes.⁸⁵

⁸⁰Rawls, *Theory of Justice*, pp. 221–228.

⁸¹*ibid.*, pp. 224–227.

⁸²*ibid.*, pp. 65–83.

⁸³*ibid.*, p. 201.

⁸⁴*ibid.*, p. 359.

⁸⁵Cohen, 'Deliberation and Democratic Legitimacy', pp. 19–20 suggests that Rawls may also

The circumstances of the constitutional stage model these considerations. That is, once deliberators are constrained to adopt a 'constitution that satisfies the two principles of justice and is best calculated to lead to just and effective legislation', they would agree on a liberal democracy.⁸⁶ Once again, if real people accept these considerations then they should also accept this conclusion. If so, then the constitutional variant of the contractualist test of legitimacy requires that uses of power by the state must be authorized by democratic means with the state's actual members participating in this decision. This means that the actual state must employ democratic institutions when making decisions. Thus concludes Rawls' justification of democratic institutions.

Even so, there remains a potential worry that this justification of democratic institutions involves an unstable hybrid. Its imperfect and pure procedural components appear in tension. In particular, Rawls grants limits on the principle of participation provided that they 'fall equally upon everyone' and they 'lead to a more just body of legislation.'⁸⁷ That is, the justice of the process may be sacrificed in order to better promote the justice of outcomes. The concern is how much sacrifice is permissible. If the computer simulation truly makes better decisions, then it seems that this account allows the disposal of all democratic processes—when the computer decides everything, each person is completely and equally disenfranchised from the political process. The epistemic aspect overwhelms.⁸⁸ A quasi-pure procedural justification of democratic institutions must carefully explain how to weigh having a just procedure against making just outcomes. The proper weight is almost certainly up for reasonable disagreement and so it is unclear what precise arrangement of democratic institutions deliberators at the constitutional stage would actually agree upon.⁸⁹

assume what I call in section 2.3.2 the mirroring strategy. According to this, democratic processes may reliably promote the two principles of justice because they are structurally similar to the original position. While the use of this strategy is not entirely obvious in Rawls' earlier work, it appears in his later work where, for instance, he claims that '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', John Rawls, 'The Basic Liberties and Their Priority', in *The Tanner Lectures on Human Values*, ed. by Sterling M. McMurrin, vol. 3 (Salt Lake City: University of Utah Press, 1982): 1–87, p. 45.

⁸⁶*Theory of Justice*, p. 197.

⁸⁷*ibid.*, pp. 228–229.

⁸⁸A quasi-pure procedural justification could also be unstable in the other direction. For instance, if the principle of participation gives no concessions to the need for making better decisions, then standards of justice for assessing outcomes carry no weight. The purely procedural aspect overwhelms, and the worry is that democratic processes can then allow decisions that manifestly violate fair equality of opportunity and the difference principle.

⁸⁹Note that Rawls cannot appeal to democratic processes for resolving this because the disagreement is precisely over their particular structure. However, his remark that 'what is of great urgency is

This is part of a more general concern with Rawls' contractualist approach: it is difficult to anticipate in advance what hypothetical agreements deliberators would make. Not only is there uncertainty about what agreements would be reached at the constitutional stage, but there is also uncertainty about whether deliberators in the original position would agree on Rawls' two principles of justice. One of the more famous debates concerns how deliberators in the original position would assess potential risks.⁹⁰ Rawls argues that they would be fairly conservative because of their uncertainty, due to the veil of ignorance, concerning their respective positions in society. This risk aversion, according to Rawls, would have deliberators adopt maximin style reasoning, which in turn would lead deliberators to agree on the two principles of justice. Critics reject the claim that uncertainty requires deliberators to reason so conservatively. These critics maintain that deliberators would attempt to maximize expected utility, leading them to adopt more utilitarian principles for governing the arrangement of the state's basic institutions. This criticism may not be devastating to Rawls' account, but it does show that reasonable disagreements like this make it difficult to anticipating the content of hypothetical agreements in the original position. This concern only increases when considering possible agreements from the later stages of the sequence.

Reasonable disagreements also concern the considerations that the original position is supposed to model. For instance, recall that one of these is that 'no one should be advantaged or disadvantaged by natural fortune or social circumstances in the choice of principles.'⁹¹ This leads Rawls to claim that the distribution of natural abilities and talents among people is arbitrary and that the state should treat it as a 'common asset' whose benefits both the talented and untalented share in a way satisfying the difference principle.⁹² Robert Nozick, along with other libertarians, rejects this, arguing that treating talents as common assets disrespects the inviolability of persons and flouts the Kantian injunction to treat persons as ends and never as means only—echoing the criticisms that Rawls makes against utilitarianism.⁹³ Nozick believes that a person should be able to use her natural

consensus on those essentials', Rawls, *Justice as Fairness: A Restatement*, p. 28, suggests that he apparently believes that unanimity is still possible at the constitutional stage given that the deliberators are still under severe informational limitations from the veil of ignorance.

⁹⁰See the debate between 'Some Reasons for the Maximin Criterion', *American Economic Review* 64, 2 (May 1974): 141–146 and John C. Harsanyi, 'Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory', *American Political Science Review* 69, 2 (June 1975): 594–606.

⁹¹Rawls, *Theory of Justice*, p. 18.

⁹²*ibid.*, p. 101.

⁹³Nozick, *Anarchy, State, and Utopia*, p. 228.

fortunes to benefit herself as she sees fit.⁹⁴ The only restriction is that she does not use them to harm others or impede their freedom. Once again, this may not be a devastating criticism, but it certainly seems to isolate another point of reasonable disagreement with Rawls' approach.

The fact of reasonable pluralism suggests that disagreements such as these may have no foreseeable resolution. General acceptance on the conditions that the original position should model or on what agreements the parties there would reach is unlikely. Unfortunately, Rawls' justification of the two principles of justice depends on such acceptance, and his quasi-pure procedural justification of liberal democracy presumes these two principles. As a result, this justification appears to have the same faults as the respective approaches of Bentham and Rousseau: it requires acceptance of a controversial comprehensive doctrine, or at least aspects of one, concerning the two principles of justice.

Rawls' later work acknowledges that the fact of reasonable pluralism potentially undermines this approach. He modifies his argument to only claim that the political processes of democratic government along with assurances of some basic liberties would be chosen as part of an idealized constitutional consensus between reasonable persons, though for reasons that may not have to do with the two principles of justice.⁹⁵ That is, deliberators at the constitutional convention are no longer constrained to agree on a constitution that both promotes and is consistent with the two principles of justice. Deliberators instead discover that their respective reasonable comprehensive doctrines admit of sufficient overlap for grounding a consensus on using democratic processes to resolve their political disagreements on how the state should exercise its power. Rawls claims that certain factors confronting the deliberators would contribute to reasonable people reaching such a consensus. That is, societal pluralism along with a desire for social stability and civil peace would cause deliberators to agree that, at the very least, democratic processes provide 'the only workable alternative to endless and destructive civil strife.'⁹⁶ Rawls does not believe that this is over idealistic or utopian, for he notes that similar circumstances lead to a consensus on principles of religious toleration following the Reformation.⁹⁷

Consequently, Rawls continues to defend a constitutional variant of the contractualist test of legitimacy for assessing state decisions while arguing that this test requires that the state adopt democratic institutions. A concern lingers with this new approach, however, is that it is still speculative concerning the outcome of the constitutional convention. At least according to the old approach, democratic

⁹⁴Nozick, *Anarchy, State, and Utopia*, pp. 167–174.

⁹⁵Rawls, *Political Liberalism*, pp. 158–164.

⁹⁶*ibid.*, p. 159. Also see my comments in section 1.2.2 (see n. 49).

⁹⁷*ibid.*, pp. xxiv–xxviii, 159.

institutions have an independent justification, though it presumes a prior agreement on the two principles of justice. The new approach, on the other hand, does not justify these institutions by appealing to a prior standard that is independent from them being agreed upon at a constitutional convention. Furthermore, Rawls only claims to anticipate that a constitutional consensus would emerge based on the reasonableness of the deliberators along with psychological and sociological contingencies currently entrenched in the contemporary culture.

To briefly summarize, this constitutional variant of the contractualist test of legitimacy attempts to justify democratic institutions by arguing that idealized deliberators would agree to include these institutions among the state's constitutional essentials. This entails that legitimate exercises of state power must be done in accordance with the processes that these institutions require. The problem with this approach is that it requires either stipulating or conjecturing that participants would agree to adopt these institutions. To answer this concern, one may argue that participants either must do so based on prior agreements that they would have made, as in Rawls' early sequential approach, or would do so based on certain unavoidable contingencies affecting the participants, as in Rawls' later approach. The first runs directly afoul of the fact of reasonable pluralism while the second remains more speculative. In response to concerns like these, deeply democratic contractualists offer an alternative approach for justifying liberal democracy.

2.3.2 Deeply Democratic Contractualism

Both Jürgen Habermas and Joshua Cohen do not believe that John Rawls' approach establishes the proper relationship between liberal democracy and legitimacy.⁹⁸ For his part, Cohen is not satisfied with the Rawlsian strategy of justifying liberal democracy by claiming that its institutions would pass the contractualist test of legitimacy. According to Cohen, this derivation is 'too indirect and instrumental' because 'it rests on a series of highly speculative sociological and psychological judgments.'⁹⁹ Cohen's approach begins with the idea that the contractualist

⁹⁸Cohen, 'Deliberation and Democratic Legitimacy', 'Procedure and Substance in Deliberative Democracy', 'Democracy and Liberty', in *Deliberative Democracy*, ed. by Jon Elster (Cambridge, MA: MIT Press, 1998): 185–231, 'Reflections on Habermas on Democracy', *Ratio Juris* 12, 4 (December 1999): 385–416; and Habermas, 'Discourse Ethics', 'Remarks on Discourse Ethics', in *Justification and Application: Remarks on Discourse Ethics*, trans. by Ciaran P. Cronin (Cambridge, MA: MIT Press, 1993): 19–111, *Between Facts and Norms, Inclusion of the Other*. I do not mean to suggest that both thinkers offer identical strategies for justifying democracy. For my purposes in this section, both contractualist approaches are sufficiently similar, though I do note a few of the places where they diverge. Regardless, Cohen and Habermas clearly display the influence they have had on each other. This is made explicit, for instance, in Cohen, 'Deliberation and Democratic Legitimacy', p. 33n.12 and Habermas, *Between Facts and Norms*, pp. 304–306.

⁹⁹Cohen, 'Deliberation and Democratic Legitimacy', p. 20.

test of legitimacy instead ought to embody 'democracy itself as a fundamental political ideal and not simply as a derivative ideal'.¹⁰⁰ That is, this test need not justify democracy by claiming that the content of hypothetical agreements, either on principles of justice or on constitutional essentials, entail the adoption of liberal democracy; the contractualist test of legitimacy instead recommends that the idealized process for making hypothetical agreements is itself explicitly democratic. Considerations like this leads Cohen to reject the Rawlsian apparatus of the original position in favor of a democratically structured contractualist device, which he calls the 'ideal deliberative procedure', whose circumstances the state should model when making actual decisions.¹⁰¹

Habermas insists on a similar democratic device—the 'ideal speech situation'—as a model for aspects of state decision making.¹⁰² His disapproval of the Rawlsian approach, however, emphasizes what he calls its 'monological form', by which Habermas means that it has the 'form of a hypothetical process of argumentation occurring in the individual mind'.¹⁰³ This occurs in two ways in Rawls' theory. First, the veil of ignorance prevents any meaningful way to differentiate the deliberators, which completely eliminates the need for them to actually deliberate with each other, for no one has any information that another lacks. Deliberation only occurs within each individual participant, not between them. As Rawls says, 'if anyone after due reflection prefers a conception of justice to another, then they all do, and a unanimous agreement can be reached'.¹⁰⁴ This leads to the second way in which Rawls' theory is monological: it allows Rawls to deliberate alone behind the veil of ignorance and stipulate unanimous agreement on his two principles of justice and the constitutional essentials of a liberal democracy. Rawls is the new philosopher-king, though a benevolent one.

According to Habermas, this monological approach leads to a problematic for an account of legitimacy by leaving actual people 'subject to principles and norms that have been anticipated in theory and have already become institutionalized beyond their control' and that 'from their perspective all of the essential discourses

¹⁰⁰Cohen, 'Deliberation and Democratic Legitimacy', p. 17.

¹⁰¹ibid., pp. 20–23.

¹⁰²This is more apparent in Habermas' earlier works, especially Jürgen Habermas, *Legitimation Crisis*, trans. by Thomas McCarthy (Boston: Beacon Press, 1975), pp. 95–143. As James Bohman, 'Complexity, Pluralism, and the Constitutional State: On Habermas's *Faktizität und Geltung*', *Law and Society Review* 28, 4 (1994): 897–930, pp. 903–906 notes, Habermas, *Between Facts and Norms* weakens the direct relationship between ideal discourse and democratic decision making in light of the complexities of political decision making and the fact of reasonable pluralism.

¹⁰³'Discourse Ethics', p. 68. Also see ibid., pp. 66–67, 94 and Jürgen Habermas, 'Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism', *Journal of Philosophy* 92, 3 (March 1995): 109–131.

¹⁰⁴Rawls, *Theory of Justice*, p. 139.

of legitimation have already taken place within the theory; and they find the results of the theory already sedimented in the constitution.¹⁰⁵ In other words, Habermas accuses Rawls of placing fundamental principles of justice and constitutional essentials beyond revision or reconsideration by those subject to them. Indeed, Habermas wants to allow deliberation on these fundamental ideas to continue. He maintains that if decisions, even hypothetical ones, are to be legitimately binding, then these decisions must have a publicly recognized connection to the outcomes of actual deliberations between actual people. As a result, Habermas advocates a 'dialogical' theory of legitimacy that denies that the philosopher may act as the expert and stipulate in advance the outcomes of hypothetical deliberation. It is only through actual deliberations modeled on the idea speech situation, according to Habermas, that garners the best evidence for determining the results of those hypothetical deliberations.

So by different routes, Cohen and Habermas reach deeply democratic conclusions, by which I mean that they hold that the contractualist test of legitimacy involves a process of hypothetical agreement that is democratic in nature. Accordingly, each has a contractualist device specifying an ideal democratic procedure. Furthermore, Cohen and Habermas offer 'deeply' democratic views insofar as they hold that *only* decisions produced in accord with this idea that are legitimately binding. Principles of justice, systems of rights and liberties, and constitutional essentials have no prior authority independent of having been chosen by an ideal democratic procedure. This does not mean that Cohen and Habermas advocate strictly majoritarian democracy, however. Both remain liberal democrats, but they hold that a system of rights is necessitated by democratic processes, for without basic liberties the ideal democratic procedure would be impossible to carry out.¹⁰⁶ Otherwise, ideal democratic deliberation cannot be constrained by any prior norms.

According to both of these approaches, decision making procedures employed by the state should mirror or model the ideal circumstances for a fair agreement. The assumption is that the similarity between actual procedures and the ideal allows one to infer that actual outcomes likely represent a legitimately binding agreement or, at the very least, provide the best evidence for what such an agreement would look like. Since both Cohen and Habermas believe that the ideal circumstances for making a legitimate agreement are inherently democratic, this mirroring strategy may then justify the adoption of a system of democratic institutions by arguing that such a system is structural similar to, and guarantees similar background circumstances, as the ideal procedure.¹⁰⁷ Therefore, they claim that

¹⁰⁵Habermas, 'Rawls's Political Liberalism', p. 130.

¹⁰⁶*Between Facts and Norms*, pp. 121–131.

¹⁰⁷Nelson, 'The Very Idea of Pure Procedural Justice', and Barry, *Justice as Impartiality*, also adopt

democratic institutions are reliable methods for making decisions in accord with the contractualist test of legitimacy.

The difference between this and the constitutional contractualism of Rawls is that the mirroring strategy need not specify in advance what fair agreements would be made by the ideal procedure in order to justify democratic institutions. The possible content of these agreements is instead revealed by the actual agreements made by real people under circumstances approximating the ideal. This means that the complaint of binding real people to hypothetical agreements has significantly less force for this strategy. The constitutive members of the state are only bound by the actual agreements they make, with the assumption again being that these agreements are sufficiently close to what the ideal would have been. So while legitimacy may be based on the adherence to hypothetical agreements, the mirroring strategy posits that actual agreements made through democratic processes are close enough to the ideal.

The mirroring strategy, however, faces some difficulties. The ideal democratic procedures Cohen and Habermas advocate make demands on the deliberators and their circumstances that are extremely difficult to satisfy in the real world. Of course, Cohen and Habermas acknowledge this, recognizing that real-world instantiations can never be exactly like the ideal. For instance, Cohen requires that deliberators have equal bargaining advantages, or at least no such advantages attached to their wealth or social status.¹⁰⁸ In reality this is never completely true: one's money and social status often do play some role in influencing political outcomes. Similarly, Habermas' ideal communication community has an unlimited time for deliberation, which is obviously unrealistic. If actual democratic institutions made demands like these, then it is doubtful that any decisions beyond the most rudimentary could be reached.

This leads to a problem that John Elster calls the 'approximation assumption'.¹⁰⁹ As economists have long noted, 'it is *not* true that a situation in which more, but not all, of the optimum conditions are fulfilled is necessarily, or is even more likely to be, superior to a situation in which fewer are fulfilled'.¹¹⁰ The concern is that any actual system of democratic institutions inevitably falls short of the ideal demo-

the mirroring strategy for justifying democratic institutions. Rawls also appears to adopt this line of reasoning when he argues that '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', Rawls, 'Basic Liberties and Their Priority', p. 45.

¹⁰⁸Cohen, 'Deliberation and Democratic Legitimacy', pp. 22–23.

¹⁰⁹This argument is presented in a different context in Jon Elster, 'The Market and the Forum: Three Varieties of Political Theory', in *The Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 103–132, p. 119.

¹¹⁰R. G. Lipsey and Kelvin Lancaster, 'The General Theory of Second Best', *Review of Economic Studies* 24, 1 (1956–1957): 11–32, p. 12.

cratic procedure outlined by Cohen and Habermas. If legitimate outcomes are just those returned by the ideal process, the extent to which the imperfect outcomes of real democratic processes are legitimate remains in doubt. No matter what improvements are made to more closely approximate the ideal, say by requiring ever more stringent forms of campaign finance reform to eliminate the bargaining advantages of the rich over legislators, the constitutive members of the state may reasonably worry that it is never enough. Since ideal processes exist only at a limit that can never be reached in practice, the mirroring strategy requires an account of why ever closer approximations of an ideal procedure yield ever closer approximations of its outcomes. I know of no such account of this for contractualist theories.¹¹¹ Without such an account, it remains a real possibility that a slight refinement, which is always possible, would yield a radically different outcomes.

A second concern with the mirroring strategy involves the reasons and justifications to which a deliberator—either ideal or actual—may appeal when rejecting or accepting a proposal. According to Cohen, the mirroring strategy requires that both ideal and actual deliberators appeal to the same sort of reasons and justifications: ‘we cannot expect outcomes that advance the common good unless people are looking for them.’¹¹² Accordingly, both ideal and actual deliberations must involve appeals to the common good.¹¹³

There is, however, an ambiguity in Cohen’s discussion of the common good. At times Cohen sounds as if it is deliberation that determines the contents of the common good. For instance, he claims that ‘the interests, aims, and ideals that comprise the common good are those that survive deliberation.’¹¹⁴ If so, then it is no longer clear what a deliberator can actually appeal to during deliberation, for he or she cannot appeal to the common good since it is not known until the resolution of deliberation. Perhaps Cohen intends actual deliberations to appeal to the common good, the contents of which are understood as whatever the ideal deliberation would produce. Doing so, though, requires abandoning the mirroring strategy, for continuing to maintain it then generates an unpleasant regress where ideal deliberators must then appeal to the outcomes of *ideal* ideal deliberations, which then necessitates the appeal to the outcomes of ideal ideal deliberations, and so on.¹¹⁵ The upshot is that if Cohen sees the common

¹¹¹One discussion of approximation issues for moral theory appears in Avishai Margalit, ‘Ideals and Second Bests’, in *Philosophy for Education*, ed. by Seymour Fox (Jerusalem: Van Leer Foundation, 1983): 77–90.

¹¹²Cohen, ‘Deliberation and Democratic Legitimacy’, p. 20.

¹¹³*ibid.*, pp. 23–25.

¹¹⁴*ibid.*, p. 25.

¹¹⁵A similar type of regress is discussed in David M. Estlund, ‘The Democracy/Contractualism Analogy’, *Philosophy and Public Affairs* 31, 4 (Autumn 2003): 387–412, pp. 400–402.

good as the product of ideal process, in typical contractualist fashion, but also wants to allow the actual processes to permit appeals to the common good, then the mirroring strategy does not seem promising.

Even so, this is not the only possible interpretation of Cohen's position, for at other times he seems to assume that the actual and ideal deliberators should appeal to a conception of the common good that is independent of the outcome of their deliberations. This is apparent when Cohen argues that the common good places constraints on the reasons that deliberators may present during the process.¹¹⁶ Assuming an independent criterion of the common good makes Cohen's position look more like that of Rousseau.¹¹⁷ If that is the case, then Cohen is reverting to a more epistemic defense of democracy, and that exposes Cohen to the problem of justifying a particular conception of the common good in the face of reasonable disagreement over its potential contents.

However, as Frank Michelman notes, the very idea of deep democracy faces a serious foundational difficulty because it rests on a worrisome regress.¹¹⁸ First, deep democracy holds that legitimate exercises of state power are those that can only meet with the assent of all subject to it after a process of democratic deliberation. Second, deep democracy holds that there cannot be prior constraints on a decision that have not also been agreed to after democratic deliberation. Third, the constraints posed by the very process of democratic deliberation must themselves have been the product of democratic deliberation. The regress begins when trying to justify a particular democratic procedure—ideal or actual—for legitimate decision making. For there must be a prior agreement outlining the use of that process. But that prior agreement itself is only justified to the extent that it too was the product of a prior democratic process. But that prior democratic process is justified because it was the product of a yet further prior agreement. And so the regress goes.

In general, the contractualist approaches of Rawls, Cohen, and Habermas are not satisfying because legitimacy itself has nothing to do with what actual people have authorized the state to do. Satisfying the contractualist test of legitimacy shows, at best, that the exercise of state power requires a certain justification (as a rational and reasonable person, you would accept this) according to their point of view, but not that it has actually been authorized by the state's constitutive members. Even Thomas Hobbes has an argument of this form, when he argues

¹¹⁶Cohen, 'Procedure and Substance in Deliberative Democracy', pp. 105–106.

¹¹⁷This connection to Rousseau is made somewhat explicit in 'An Epistemic Conception of Democracy'.

¹¹⁸Frank I. Michelman, 'Review of Jürgen Habermas' *Between Facts and Norms*', *Journal of Philosophy* 93, 6 (June 1996): 307–315, and 'How Can the People Ever Make the Laws? A Critique of Deliberative Democracy'.

that the exercises of power by the sovereign ought to be justified in terms that a rational person would accept, and this hardly leads to democratic outcomes. As a result, many theories abandon contractualist approaches in favor of purely procedural ones, which dispense with the reference to hypothetical agreements. Legitimacy is found in actual decision made under real circumstances, though these circumstances must satisfy certain standards.

2.4 Purely Procedural Liberal Democracy

In response to these challenges faced by epistemic and contractualist justifications of democracy, many liberals turn towards more purely procedural ones.¹¹⁹ With persistent disagreements concerning epistemic standards or the outcome of ideal deliberation, purely procedural justifications holds that agreement can be reached on fair processes for determining when to authorize state action. Fair processes

¹¹⁹I have in mind a wide array of different theorists here, such as Joseph A. Schumpeter, *Capitalism, Socialism and Democracy*, 3rd ed. (New York: Harper and Row, 1950), part IV; Robert A. Dahl, 'Procedural Democracy', in *Philosophy, Politics, and Society: Fifth Series*, ed. by Peter Laslett and James S. Fishkin (New Haven, CT: Yale University Press, 1979): 97–133, and *Democracy and Its Critics*; Thomas Christiano, 'Freedom, Consensus and Equality in Collective Decision Making', *Ethics* 101, 1 (October 1990): 151–181, 'Social Choice and Democracy', in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 173–195, *Rule of the Many: Fundamental Issues in Democratic Theory*, (Boulder, CO: Westview Press, 1996), and 'The Significance of Public Deliberation', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 243–277; Elster, 'The Market and the Forum: Three Varieties of Political Theory'; Bernard Manin, 'On Legitimacy and Political Deliberation', *Political Theory* 15, 3 (August 1987): 338–368; Cass Sunstein, 'Beyond the Republican Revival', *Yale Law Journal* 97, 8 (July 1988): 1539–1590; Stuart Hampshire, *Innocence and Experience* (Cambridge, MA: Harvard University Press, 1989); James S. Fishkin, *Democracy and Deliberation: New Directions for Democratic Reform* (New Haven, CT: Yale University Press, 1991); Jane Mansbridge, 'A Deliberative Theory of Interest Representation', in *The Politics of Interest*, ed. by Mark Petracca (Boulder, CO: Westview Press, 1992): 32–57; Norman Daniels, 'Rationing Fairly: Programmatic Considerations', *Bioethics* 7, 2–3 (1993): 224–233, and 'Accountability for Reasonableness: Establishing a Fair Process for Priority Setting is Easier than Agreeing on Principles', *British Medical Journal* 321, 7272 (November 2000): 1300–1301; Norman Daniels and James Sabin, 'Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers', *Philosophy and Public Affairs* 26, 4 (Autumn 1997): 303–350, 'Last Chance Therapies and Managed Care: Pluralism, Fair Procedures, and Legitimacy', *Hastings Center Report* 28, 2 (March 1998): 27–41, and *Setting Limits Fairly: Can We Learn to Share Medical Resources?*; Jack Knight and James Johnson, 'Aggregation and Deliberation: On the Possibility of Democratic Legitimacy', *Political Theory* 22, 2 (May 1994): 277–296; Seyla Benhabib, 'Toward a Deliberative Model of Democratic Legitimacy', in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. by Seyla Benhabib (Princeton: Princeton University Press, 1996): 67–94; and John S. Dryzek, *Deliberative Democracy and Beyond: Liberals, Critics, Contestations* (Oxford: Oxford University Press, 2000).

allow participants to weigh their competing values and principles and create working solutions. So while the liberal state may be divided on controversial issues, legitimately binding decisions can still be made, provided that they have the appropriate procedural pedigree and regardless of whether some disagree with a particular outcome.

Certainly it stands in the favor of a decision that it was the product of a recognizably fair process. Nonetheless, many purely procedural democrats are not clear on why, if the disagreement over principles for justifying instances of state action is so decisive, there is greater opportunity for agreement on a fair process for doing so. This assumption seems based less on a general argument, but primarily on a failed search for appropriate political principles, a search impeded by the fact of reasonable pluralism. Nevertheless, they could probably derive it by general argument by appealing to a fairly common tradition of political thought that maintains that the fact of pluralism prevents agreement on particular social policies but not on democratic procedural institutions that choose those policies.¹²⁰

Stuart Hampshire provides a paradigmatic defense of this position.¹²¹ He maintains that all reasonable comprehensive doctrines must presuppose—and not merely acknowledge—the evil in ‘murder and the destruction of life, imprisonment, enslavement, starvation, poverty, physical pain and torture, homelessness, [and] friendlessness’.¹²² Furthermore, all reasonable views should agree that fair processes of negotiation are necessary to prevent these evils. However, there will remain those not satisfied by the decision ultimately made. Hampshire argues that this latter problem entails some basic requirements that a process must satisfy in order to be fair: a fair process should give all sides equal access, the opportunity to be heard, and the possibility to influence its outcome. A properly organized system of democratic processes best satisfies these important demands.

Other purely procedural democrats usually adopt an approach like this by spe-

¹²⁰On the descriptive end, some theorists of constitutional design maintain that the United States Constitution primarily provides rights, principles, and standards that ensure fair democratic processes. Other principles, such as those for fair distribution, are left for these processes to determine. Advocates of this reading of the Constitution include Holmes, ‘Dissenting Opinion’, in *Lockner v. New York*, 198 U.S. 45, 74 (1905), John Hart Ely, *Democracy and Distrust* (Cambridge, MA: Harvard University Press, 1980), and Bruce Ackerman, *We the People: Foundations* (Cambridge, MA: Belknap Press of Harvard University Press, 1991). While on the normative end, some political theorists argue that agreement over principles such as equal consideration lead to agreement on democratic process but do not extend further. Advocates of this position include Robert A. Dahl, *A Preface to Democratic Theory* (Chicago: University of Chicago Press, 1956), and *Democracy and Its Critics*, and Hampshire, *Innocence and Experience*.

¹²¹The following argument roughly follows *ibid.*, pp. 81–110.

¹²²*ibid.*, p. 90.

cifying similar conditions that a fair process must satisfy in order for its decisions to be legitimately binding. Such conditions include features like fair consideration and accountability for reasonableness. Fair consideration demands that a procedure ‘foster thorough deliberation about the facts, reasons, and principles that are relevant to the dispute’.¹²³ This is perhaps trivial, though it is certainly an important baseline requirement. Accountable for reasonableness is far more demanding. This is composed of four different conditions. First, it requires that the decisions of a process and their rationale be publicly available (publicity). Second, the rationales for these decisions should appeal to ‘evidence, reasons, and principles that are accepted as relevant by fair-minded people who are disposed to finding mutually justifiable terms of cooperation’ (relevance).¹²⁴ Third, the institutions must provide means for stakeholders to challenge the process’ outcomes, allowing for their revision in response to new evidence or arguments (revision and appeals). Finally, there must be voluntary or public regulation to ensure the previous conditions are satisfied (regulation). Purely procedural democrats believe that conditions like these may further fair consideration as well as answer the challenge of legitimacy: those dissatisfied with an outcome can nonetheless accept it because their judgments were duly considered and it was not arbitrarily chosen but justified by reasoning they can fairly appreciate.¹²⁵

Fair consideration and accountability for reasonableness may then enumerate some relatively straightforward conditions that a fair process should satisfy. First and foremost, fair consideration demands *minimal fairness*, where a process must guard the state’s constitutive members (and perhaps even non-members) against unacceptable forms of power and dominance. A fair process cannot exclude some people or favor some inappropriately. This entails what I hope are two rather obvious criteria for a fair procedure.

To begin with, minimal fairness demands that a fair process cannot simply ignore the judgments of *all* the stakeholders when evaluating the alternatives. For instance, a process cannot be rigged to always favor one view over all others. Even if this judgment is expressible in terms that everyone can fairly accept as relevant, this is clearly a terrible process. To take an extreme case, suppose that deliberation lead everyone to agree that one course of action is the one to pursue, but a process

¹²³ Daniels and Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, p. 4.

¹²⁴ *ibid.*, p. 45.

¹²⁵ But consider the following case: a ‘reasonable dictator’ could make her decisions public and publicize her rationales, which are couched in terms that reasonable people can accept. She also provides a venue for appeals, kindly listens as the other stakeholders present their views, and voluntarily regulates her decisions. Despite all this, however, she ignores all other rationales, and consistently applies only her values to the situation. This seems consistent with accountability for reasonableness, but it hardly amounts to fair consideration.

returns the opposite. This process simply imposes its decision, irrespective of the precise contents of the individuals' evaluations. As such, this process is hardly fair since it has simply disregarded everyone's input. A fair process must therefore satisfy a condition of *non-imposition*. Specifically, a process cannot return the same judgments irrespective of whatever evaluations the participants might possess.

Another concern is that the process would only consider the evaluations of a single stakeholder, ignoring those of the others, regardless of what their evaluations may be. Minimal fairness also prohibits this. The President of the United State cannot simply do whatever she wants based solely on her own personal convictions. Even if her position is accountable for its reasonableness, such a dictatorship removes all but that individual from consideration and consequently this process is not fair. The President may remain the ultimate decision maker, but her decisions must demonstrate consideration of the views of the others affected by her decisions. This need is explicit in the requirement for processes of revision and appeals. These processes are supposed to ensure that everyone's evaluations receive a hearing and are fairly incorporated into the decision or revision. A fair procedure must then satisfy *non-dictatorship*: decisions cannot be based solely on the evaluations of one person, irrespective of what the other people might hold.

Fair consideration also demands *formal fairness*. This requires that like cases are treated alike. Of course, the crucial challenge is determining what makes like cases alike. According to one criterion, two cases are alike for a given set of alternatives if all the state's constitutive members' respective evaluations concerning those options are identical in both cases. With this criterion, formal fairness demands that a process select the same option every time it is given the same individual judgments. Certainly as the alternatives change or individual evaluations of them change, the process may choose differently, but when they do not, the process must remain consistent. This condition has immediate impact. For instance, it bars processes that randomly select a dictator who then chooses, and it denies processes from randomly making the decision. In both cases, the very same profile of individual evaluations can lead to radically different outcomes. It is the arbitrary fashion in which such stochastic processes decide that makes them unfair.

The primary consequence of this interpretation of formal fairness is that it entails that the only relevant information for making a fair decision lies in the evaluations of the state's constitutive members. These people would be rightfully suspicious of outcomes based in part on information—even when it is accountable for its reasonableness—that is independent of all of their evaluations for the same reason they can rightfully deny the imposed outcomes from before. They would ask, 'Where did this independent information come from? If it affects none of our evaluations of the alternatives, why is it relevant?' The onus is then on those denying this condition to answer these questions.

Another condition that fair consideration entails is a condition of *unanimity*. This condition requires that when the evaluations of all stakeholders unanimously agree that one alternative is superior to another, the process' judgments also agree with this. To clarify, this condition applies to partial consensuses at the level of evaluations and not necessarily to those at the level of principles. So whenever there is such a partial consensus concerning some alternatives, the process should preserve this, just as it should respect a (highly improbable) consensus concerning all the alternatives. If this condition of unanimity did not hold, stakeholders would rightfully demand to know what authority the process has to overturn their collective agreement.

The final condition entailed by fair consideration that I would like to consider is *unrestricted admissibility*. This condition requires a process to be prepared to handle all the logically possible evaluations that a stakeholder might ultimately possess. This may sound like more a formal concern than one of fairness, but this is still a requirement of fair consideration. Processes that violate this condition would bar otherwise reasonable evaluations from consideration by imposing a class of 'admissible' evaluations from which each stakeholder may choose. Stakeholders would have reason for concern as to how this delineation was determined or who chose it. Outcomes from such a process would be especially illegitimate to those who have one of the 'barred' evaluations and are capable of defending it in deliberation.

In light of all these considerations, the most commonly invoked fair process for democracy is voting, as it possesses some appealing features. For one thing, voting appears to ensure impartial and equal treatment of society's members by giving each person an equal vote and an equal chance to influence the outcome. Voting can also treat the alternatives in a neutral manner. That is, the voting process can be constructed so that it does not provide an advantage to any particular option. In addition, voting is supposed to obviate the need for society to endorse the truth of any of the competing reasonable doctrines, unless an overwhelming majority accepts it. According to this account, each person is free to evaluate the alternatives, for, after all, mature adults are thought to be the best judges of what satisfies their own values, while society is still able to enact policies by employing the sanctioned voting process. Outcomes are then binding because they were chosen in a manner responsive to the competing values and judgments without imposing any on society's members.

The application of voting processes to resolve social disputes seems fairly straightforward. The state proposes various social policies and its stakeholders simply vote. Some purely procedural democrats, however, argue that voting is not

sufficient in such circumstances.¹²⁶ In typical voting scenarios, individuals present their evaluations and the process selects an option that satisfies the most of these evaluations. The obvious concern here is tyranny of the majority. This might not be an objection when a committee is selecting what to order for dinner on a Friday night, but it is serious when resolving fundamental moral disputes concerning issues such as healthcare, abortion, and business regulation. A person who prefers fish and chips, but is otherwise perfectly capable of enjoying a hamburger, is not tyrannized when the majority determines that hamburgers are ordered. Things look differently, however, when the person favors fish and chips because of Lenten dietary restrictions or an allergenic reaction; it is somewhat tyrannical for the majority to disregard this and still order hamburgers.

This highlights what a fundamental difference between conflicts in matters of tastes and conflicts in matters of moral values and judgments. According to most purely procedural democrats, processes for addressing the latter types of conflicts must involve fair consideration and accountability for reasonableness. Addressing moral conflicts is typically thought to require thorough deliberation and argumentation, where individuals present their evaluations and the rationale behind them. In these deliberations, the reasons should be convincing to, or at least be seen as relevant by, all the interlocutors. Ultimately the deciding factor is the force of the better argument, not the number of people shouting it. All this places constraints on a process that voting does not necessarily satisfy, as voting only requires individuals to present their evaluations, nothing further. Voting ensures a minimum of fair consideration, but limited accountability for reasonableness. As a result, using majority rule to navigate moral conflicts provides a rather weak thread binding those dissatisfied with the outcome, for they are only informed that a majority of people believe otherwise. This thread is plausibly sufficient when matters of taste are involved, but not with contesting moral convictions. A person against whom the government is willing to employ its coercive power, say through increased taxes, is morally entitled to know the reasons for why this is so, and these reasons must be in terms that person recognizes as relevant. The result of a vote is often not enough.

Nevertheless, most purely procedural democrats do not dismiss voting entirely. Recalling the fact of reasonable pluralism, it is highly unrealistic that argumentation and deliberation will reliably produce a consensus concerning controversial social issues. Further processes are then necessary to determine the final outcome. Therefore, an integrated process is usually suggested, with a period of deliberation followed by a vote when necessary. Deliberation should be structured, however,

¹²⁶This and the next paragraph roughly follow Daniels and Sabin, *Setting Limits Fairly: Can We Learn to Share Medical Resources?*, pp. 34–36.

to provide constraints on the reasons that stakeholders may appeal to in justifying their evaluations. Once this is done, and if disagreements nevertheless remain, then a vote settles the issue. The decisions of such a process are supposed to have binding force because the rationale justifying this decision—and not merely the procedural facts that led to it—become transparent and it is in terms that all stakeholders can accept as relevant. Therefore, accountability for reasonableness is thought to seriously dampen the threat of a tyrannical majority because, for example, denying the fundamental liberties of a minority could hardly be justified in terms they could or should recognize. Coupling public deliberation with voting processes is therefore intended to ensure more robust satisfaction of fair consideration and accountability for reasonableness.

It is worth appending to this that some purely procedural democrats do not require all fair processes to incorporate voting.¹²⁷ This does not obviate the liberal state's duty to fair consideration and accountability for reasonableness in its decisions, however. As with those resulting from voting processes, legitimately binding decisions require deliberative processes and must be justifiable by reasons all stakeholders accept as relevant. This demand applies even if all the state's constitutive members cannot directly participate in the decision-making process. So all the liberal state's institutions do not need to hold elections, but they need some transparent mechanism for navigating the conflicts between the reasonable views of its members.

As this caveat illustrates, many different processes may satisfy fair consideration and accountability for reasonableness, which most purely procedural democrats recognize.¹²⁸ Undoubtedly there are many ways to structure agenda formation, the procedure for introducing alternatives, the deliberations over these alternatives, and the final selection of an alternative when conflicts among stakeholder evaluations remain. Fair consideration and accountability certainly limit the processes that the liberal state may employ in handling these matters, but it still retains some flexibility. A prominent concern, though, remains for how the liberal state ought to choose between different fair processes. Consider process *X*, which consists of public deliberation coupled with majority rule, and process *Y*, which consists of public deliberation and expert consultation but an executive who ultimately makes the decision. If a minority believes that the state adopts a course of action that is unacceptable oppressive to its members under process *X* but gets an outcome its members favor with process *Y*, then how the liberal state chooses between adopt-

¹²⁷Dryzek, *Deliberative Democracy and Beyond: Liberals, Critics, Contestations*, pp. 50–55, is a notable advocate of this position. He proposes that actions like protests and boycotts in civil society provide an alternative to voting.

¹²⁸Daniels and Sabin, 'Last Chance Therapies and Managed Care: Pluralism, Fair Procedures, and Legitimacy', 38ff.

ing X and Y is of utmost importance to this group of people. With a plurality of fair processes to select from, this decision begins to look as contentious as that over fair principles. The flight to pure proceduralism may not get us far.

The majority of work by purely procedural democrats is to address this issue. However, instead of pursuing this concern, I would rather consider the converse problem, which is often taken for granted or simply ignored. Given the demands that fair consideration and accountability for reasonableness make on a fair process, formal results in social choice theory suggest that there is no fair process to choose. According to these results, choosing a process inevitably involves giving up some component of fairness or some basic principle of rational decision making. The implication is then that liberal democracy faces serious problems to the extent that democratic institutions are justified on the basis of their fairness. Such a concern has punch for purely procedural democrats, but also on epistemic democrats who also appeal to the fairness of democracy. I now wish to turn to these results in order to assess these implications.

CHAPTER THREE

Social Choice Theory and Liberal Democracy

DEMOCRATIC institutions, regardless of whether they are understood epistemically or purely procedurally, attempt to provide a way for the state to make legitimately binding decisions in the face of societal pluralism. One of the earliest known accounts specifying formal procedures to follow when making such collective decisions appears in Ramon Llull's novel *Blanquerna*.¹ Speaking through the character Sister Natana, Llull prescribed a procedure for a convent of nuns to elect their mother superior. Following a vow of honesty, each sister was asked to vote over every pair of candidates; the candidate with the most votes across all these comparisons would be elected. With a field of nine candidates, a single ballot then consisted of thirty-six pairwise comparisons. 'By this method', Natana maintained, 'the truth is found'.

Five centuries later during the French Revolution, the spirit of the European Enlightenment spurred many of the revolutionaries to seek methods for the reasoned construction of social order and collective decision making that took into account the judgments of every member of society. One French mathematician, the Marquis de Condorcet, rediscovered the pairwise approach to decision making,

¹A translation of an excerpt, 'From *Blanquerna*, Chapter 24, ca. 1283', may be found in Iain S. McLean and Arnold Urken, eds., *Classics of Social Choice*, trans. by Iain S. McLean and Arnold Urken (Ann Arbor: University of Michigan Press, 1995), pp. 71–73. Günter Hägele and Friedrich Pukelsheim, 'Llull's Writings on Electoral Systems', *Studia Lulliana* 41, 97 (2001): 3–38 provide an exhaustive analysis of all Llull's known writings on electoral processes. An even earlier discussion appears in Pliny the Younger's 'Letter to Titius Aristo, A.D. 105', also in McLean and Urken, eds., trans., *Classics of Social Choice*, pp. 67–70.

though his own system suffered from the problem that it sometimes returned no winner.² At the same time, Jean-Charles de Borda, Condorcet's compatriot and rival, developed his own method, with its own faults.³ So just as the optimism of the revolution quickly turned to terror, so too did Condorcet and Borda's work fail to yield the hopeful results they desired.

In the following century and a half, great thinkers such as C. L. Dodgson (better known as Lewis Carroll) continually stumbled upon problems with social decision-making schemes.⁴ However, it was Kenneth Arrow's famous theorem that finally appeared to demonstrate that all procedures for collective decision making—even Sister Natana's—are unavoidably flawed.⁵ Informally, Arrow's theorem says that even if each individual in a group rationally orders the alternatives there is no method for fairly taking these individual judgments and returning a single, collective assessment that is also rational. The impact of this result has been spectacular: many economists, political scientists, and even philosophers now hold that collective decision making produces only arbitrariness and instability, though at least as many have proposed strategies to escape from that dismal conclusion.

In this chapter, I seek to explain why the impossibility theorems of social choice have relevance for liberal democracy, and why I believe it is particularly devastating to purely procedural understandings of democracy. To this end, I first summarize the conception of rationality presumed by these theorems (section 3.1). I then present two formal results within social choice theory—Arrow's theorem and a theorem due to Teddy Seidenfeld, Joseph Kadane, and Mark Schervish—that provide seemingly compelling evidence that it is impossible for the liberal state to make rational decisions that are responsively impartial (sections 3.2 and 3.3). Then for the remainder of this chapter, I argue why advocates of liberal democracy must take these results seriously by showing how common responses to these im-

²Marquis de Condorcet, *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris: L'Imprimerie Royale, 1785). A translation of the relevant section of this may be found as 'An Essay on the Application of Probability Theory to Plurality Decision-Making (1785) (extract: Preliminary Discussion. Analysis of Part One. Fourth Example: An Election Between Three Candidates)' in Iain S. McLean and Fiona Hewitt, eds., *Condorcet: Foundations of Social Choice and Political Theory*, trans. by Iain S. McLean and Fiona Hewitt (Brookfield, VT: Edward Elgar, 1994), pp. 120–130.

³Jean-Charles de Borda, 'Mémoire sur les Élections au Scrutin', in *Histoire de l'Académie Royale des Sciences* (Paris 1781). Fully translated as *On Elections by Ballot* in McLean and Urken, eds., trans., *Classics of Social Choice*, pp. 83–89.

⁴Charles L. Dodgson, *Facts, Figures, and Fancies, Relating to the Elections to the Hebdomal Council, the Offer of the Clarendon Trustees, and the Proposal to Convert the Parks into Cricket Grounds* (Oxford: Parker, 1874) and *The Principles of Parliamentary Representation*, (London: Harrison and Sons, 1884). These, and other works by Dodgson may be found in McLean and Urken, eds., trans., *Classics of Social Choice*, pp. 279–320.

⁵Kenneth J. Arrow, *Social Choice and Individual Values*, 2nd ed. (New York: Wiley, 1963).

possibility results either fail to escape, or simply acquiesce to, its dire implications (sections 3.4–3.7). This suggests a more radical break with the assumptions behind the impossibility results than is normally taken, which I lay out in the next chapter.

3.1 The Default Rationality Thesis

Standard economic theory assumes that a rational decision maker, call it i , possesses a value ordering R_i over a set of m potential options, denoted by the set $O = \{x, y, z, \dots\}$.⁶ Economists usually call this evaluation a ‘preference ordering’. This phrase is unfortunate, however, as it evokes egotistical or subjective connotations of anticipated satisfaction and desire. While many economists may approve of that, the formal apparatus of economic theory itself does not demand a *feeling* of preferring on the part of a rational agent. Instead, the theory permits an agent to employ whatever values, commitments, or considered judgments in order to construct an all-things-considered ranking: ‘it is simply assumed that the individual orders all [options] by whatever standards he deems relevant.’⁷ I therefore stick to a more neutral term, referring to the more general concept of a value ordering.

A value ordering is represented as a binary relation of value judgments between potential options. That is, given a value ordering R_i , xR_iy expresses the judgment that ‘according to value ordering R_i , option x is better than or equally valuable as option y ’. Based on this understanding, R_i separates into two disjuncts: superiority and indifference. Superiority is denoted by P_i , and xP_iy holds whenever xR_iy holds but yR_ix does not. Therefore, xP_iy expresses the value judgment that ‘according to R_i , option x is better than option y ’. Indifference is denoted by I_i , and xI_iy holds whenever both xR_iy and yR_ix hold. xI_iy then means that ‘according to R_i , option x is equally valuable as option y ’. Based on these definitions, superiority and indifference are mutually exclusive value judgments: given any two options, both superiority and indifference cannot hold between them. These definitions also imply that superiority and indifference are mutually exhaustive: if xR_iy holds then either xP_iy or xI_iy holds as well.

According to standard economic theory, deliberating agent i is rational provided that her value judgments are organized into a weak ordering. This means that R_i satisfies the following three conditions:

- (1) R_i is complete: for all options x and y in O , either xR_iy or yR_ix (or both) hold.

⁶Throughout this discussion, I am assuming that there is a finite number of potential options.

⁷Arrow, *Social Choice and Individual Values*, p. 17.

	xP_iy	yP_ix	xI_iy
1 st	x	y	$x \sim y$
2 nd	y	x	

FIGURE 3.1: The three ways that two options, x and y , may be ordered ($\sim$ denotes indifference between options).

- (2) R_i is reflexive: for all options x in O , xR_ix holds,
- (3) R_i is transitive: for all options x, y , and z in O , if both xR_iy and yR_iz hold, then xR_iz also holds, and

Completeness entails that, given any pair of options, the value ordering includes a judgment of them. The deliberating individual must ultimately ‘connect’ them by rendering judgment, even if it is a matter of indifference—the value ordering cannot remain agnostic. Reflexivity actually follows immediately from completeness, but it serves as a useful reminder that a single option is connected to itself. Transitivity demands a connectivity of its own: the connections between options x and y and between options y and z may entail a connection between x and z .

Requiring the value judgments in R_i to be organized into a weak ordering means that, given any two options x and y , individual i must judge that exactly one of the following must hold: xP_iy , yP_ix , or xI_iy . These three orderings are depicted pictorially in figure 3.1. In addition, the definitions of P_i and I_i , entail a whole catalogue of other requirements that must hold for them, the following five being representative:

- (1) I_i is transitive,
- (2) I_i is reflexive,
- (3) P_i is transitive,
- (4) P_i is anti-reflexive: for all options x in O , xP_ix cannot hold, and
- (5) P_i is asymmetric: for all options x and y in O , if xP_iy holds then yP_ix cannot hold.⁸

The anti-reflexivity of superiority reflects the intuition that it makes no sense to claim that an option is better than itself. Similarly, superiority’s asymmetry denies the sensibility of claiming that an option is simultaneously better and worse than another option.

Alone, a value ordering R_i does not contain any *cardinal*, or numerical, information concerning i ’s judgments about the options. It can show that one option is

⁸The proofs of these are trivial and may be found in Arrow, *Social Choice and Individual Values*, pp. 14–15. Robert Nozick, *The Normative Theory of Individual Choice* (New York: Garland, 1990), p. 85 regards these as ‘conditions of sanity’. For applications, see Hans G. Herzberger, ‘Ordinal Preference and Rational Choice’, *Econometrica* 41, 2 (March 1973): 187–237, pp. 192–196.

	v_1	v_2
x	1.0	1.0
y	0.9	0.1
z	0.0	0.0

FIGURE 3.2: Two value functions ordering three options in the same way, but expressing differing assessments for the relative value of option y .

superior to another—this *ordinal* information is contained in R_i —but it does not express the magnitude or intensity of this superiority. Many economists believe that an agent’s evaluations of options should also include this additional information.⁹ This can be done by supplanting a value ordering R_i with a value function v_i , where, for each option x , $v_i(x)$ is a real number. Naturally, the higher the value of $v_i(x)$, the better x is according to v_i . Economists usually refer to these functions as ‘utility functions’, but I avoid this term because this raises the specter of utilitarianism, which the formal structure of these functions does not necessarily demand. A value function v_i is easily converted to a weak ordering R_i : for options x and y , xR_iy holds if $v_i(x) \geq v_i(y)$. Figure 3.2 illustrates how two value functions v_1 and v_2 may order three options the same, but suggest some different assessments of the options’ relative value. So, for instance, the assessments of v_1 suggest that option y is nearly as good as option x , whereas according to v_2 , y is nearly as bad as z .

Value judgments that conform to the structure of a weak ordering are appealing because it is then clear how to determine which option(s) those evaluations regard as the best. For when presented with an agenda of available options A , which is a subset of the potential options in O , a value ordering R_i clearly indicates the best options are its top-ranked ones.¹⁰ These optimal options are denoted by the set $C(A, R_i)$, where each option in $C(A, R_i)$ is superior to the available options not in $C(A, R_i)$. Therefore, $C(A, R_i)$ is defined as follows:

⁹For instance, see John C. Harsanyi, ‘Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking’, *Journal of Political Economy* 61, 5 (October 1953): 434–435, and ‘Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility’, *Journal of Political Economy* 63, 4 (August 1955): 309–321; and Donald Davidson, J. C. C. McKinsey and Patrick Suppes, ‘Outlines of a Formal Theory of Value, I’, *Philosophy of Science* 22, 2 (April 1955): 140–160, pp. 148–157.

¹⁰It may be worth dwelling on the distinction between potential and available options for a moment. For instance, any person satisfying certain age and residency requirements is a *potential* candidate for President of the United State of America. However, only a small subset of these people, say those endorsed by a major political party, is an *available* candidate voters seriously consider for the office. So while value judgments range over all the potential options, at the time of choice, it is only those judgments concerning the available options that matter. This idea is extremely important for understanding the motivation for Independence of Irrelevant Alternatives, which I discuss in sections 3.2 and 3.6.

Optimization: $C(A, R_i) = \{x \mid x \in A \text{ and for all } y \in A, xR_i y \text{ holds}\}$.¹¹

That is to say that $C(A, R_i)$ consists of the available options from A that are top-ranked according to the value judgments in R_i . If only one optimal option may be chosen, another criteria, apart from R_i , is necessary to resolve any ties. Regardless, most economists believe that optimization of a weak ordering is the most innocuous notion of rationality. A person only needs to make decisions based on her all-things-considered ordering of the options. It is left to the individual's own values, commitments, and considered judgments to generate that ranking.

Given its apparent parsimony and perhaps even naturalness, standard economic theory holds that rational choice consists of, at the very least, the optimization of a weak ordering. I will call this the default rationality thesis. This thesis is often justified by appeal to the fact that if an agent's judgments are not consistent with a weak ordering, there may be no optimal option for that agent to choose. To illustrate, suppose Jones is an art collector and that she has the following cyclic assessments of paintings: ranking Vincent van Gogh's *Starry Night* over Salvador Dali's *Persistence of Memory*, the Dali over Claude Monet's *Water Lilies*, and the Monet over the van Gogh. According to most economists, together these judgments are irrational. Jones ranks the Monet over the van Gogh, and yet transitivity of superiority requires that she rank them the other way. Since Jones violates the asymmetry of superiority, no weak ordering can represent her evaluations. The trouble with this is that when selecting her favorite painting there is no optimal choice.¹² No matter which one she selects, there is another painting that she judges as better.

This point is typically reinforced by appeal to so-called 'money pump' arguments.¹³ Assume Jones rejects the transitivity requirement for her assessment of paintings. Now suppose she meets a devious art dealer who sees Jones' plight and so he gives her the van Gogh free of charge. Before she has a chance to run off with it, however, he then offers to give Jones the Monet in exchange for the van

¹¹This is Definition 3 in Arrow, *Social Choice and Individual Values*, p. 15. This form of optimization is similar to, but not identical with, the concept of 'optimization' as used by economists. For more discussion on optimization and its contrast with another choice rule, 'maximization', see Amartya K. Sen, 'Non-Binary Choice and Preference', in *Logic, Methodology and Philosophy of Science IX*, ed. by Dag Prawitz, Brian Skyrms and Dag Westerstaahl (Amsterdam: Elsevier Science, 1994), 'Maximization and the Act of Choice', *Econometrica* 65, 4 (July 1997): 745–779, and section 5.1.

¹²Lemma 1*1 in *Collective Choice and Social Welfare*, (Amsterdam: North-Holland, 1970), p. 16 demonstrates how cyclic judgments may entail that there are no optimal options. Of course, if a value ordering R_i is generated from a value function v_i , then R_i is guaranteed to not be cyclic.

¹³The money pump argument is attributed to Norman Dalkey by Davidson, McKimsey and Suppes, 'Outlines of a Formal Theory of Value, I', p. 146. For the standard expositions of this argument, see *ibid.*, pp. 145–146 and Howard Raiffa, *Decision Analysis: Introductory Lectures on Choice under Uncertainty* (Reading, MA: Addison-Wesley, 1968), p. 78.

Gogh along with a relatively small ‘transaction’ fee.¹⁴ By Jones’ lights, the Monet is better than the van Gogh, so her choice is simple: she accepts the offer. But then the con artist offers her the Dali for another small fee and the return of the Monet. Once more, Jones accepts this offer because she judges the Dali superior to the Monet. Next, he offers Jones the van Gogh in exchange for a small fee and the return of the Dali. And yet again, based on Jones’ evaluations, she should accept.

Following these transactions, Jones is back where she started, but with her pockets slightly emptier. However, if her assessments remain the same, the hustler is free to repeat this process until he has robbed Jones of all her disposable income or Jones’ assessments finally change—experience teaching her that transitivity is a rather good thing to obey. Money pump arguments of this sort are supposed to demonstrate that, according to the agent’s very own interests—whether they are for money or whatever else—it is necessary to structure its judgments into a weak ordering.

These arguments place the onus on those who would deny that rationality demands optimization of a weak ordering. Even so, the default rationality thesis has met little resistance, being accepted in some form by most economists and political scientists. Many ethical theorists, to the extent that they are clear on the matter, do not deny it either.¹⁵ I believe this is because the default rationality thesis has clear connections to the idea of systematicity from section 1.2. A reasonable doctrine is often thought to require a weighing and ordering of considerations

¹⁴The fee need not be in monetary units, but in units of anything that Jones happen to care about. The point, as is soon seen, is that when an agent has cyclic, intransitive preferences, the hustler is free to pump Jones of whatever she finds valuable until she none of it left with which she can willingly part.

¹⁵John Broome, *Weighing Goods: Equality, Uncertainty and Time* (Oxford: Blackwell, 1991), *Weighing Lives*, (Oxford: Oxford University Press, 2004) and Donald Regan, ‘Value, Comparability, and Choice’, in *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997): 129–150 are the most explicit ethical theorists to accept the weak ordering requirement for rationality. Even many theorists who advocate value pluralism seem to suggest that something like a single, overarching, all-things-considered, weak ordering is required for rational choice amongst options; see, for instance, Charles Taylor, ‘The Diversity of Goods’, in *Utilitarianism and Beyond*, ed. by Bernard Williams and Amartya K. Sen (Cambridge: Cambridge University Press, 1982): 129–144; ‘Leading a Life’, in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 170–183; James Griffin, *Well-Being: Its Meaning, Measurement and Moral Importance* (Oxford: Clarendon Press, 1986), ‘Incommensurability: What’s the Problem?’, in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 35–51; Michael Stocker, *Plural and Conflicting Values* (Oxford: Clarendon Press, 1990); and Ruth Chang, ‘Introduction’, in *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997): 1–34, ‘Putting Together Morality and Well-Being’, in *Practical Conflicts: New Philosophical Essays*, ed. by Peter Baumann and Monika Betzler (Cambridge: Cambridge University Press, 2004): 118–158.

to make all-things-considered judgments. The default rationality thesis simply says how to decide given such judgments. However, the impossibility results of social choice theory may be understood as challenging the idea that the state may reason rationally, as dictated by the default rationality thesis, while remaining responsively impartial to the reasonable views of its constitutive members. I now turn to two of these results.

3.2 Social Choice Theory and Arrow's Theorem

Broadly speaking, social choice theory considers issues concerning how groups ought to make decisions. In part, it examines how the state may make rational judgments while remaining responsive to the wide variety of values and commitments endorsed by its constitutive members. While Plato and Aristotle certainly consider this problem in the *Republic* and the *Politics*, it was over two millennia later when Kenneth Arrow presented a formal result appearing to demonstrate that this problem admits of no acceptable solution. In particular, his result suggests that it is impossible for the liberal state to make responsively impartial decisions that also satisfy the default rationality thesis. In fact, this theorem strongly suggests that *any process whatsoever*, whether employed by an individual or the state, for taking a set of value orderings and returning a single, all-things-considered judgment cannot be both rational and responsively impartial.

To understand what the theorem actually shows, some further formal terminology is necessary. According to Arrow, the challenge of collective decision making involves a group of n persons, denoted by the set $N = \{1, 2, \dots, n\}$, making a decision that is binding to them all. In particular, Arrow is concerned with how these people should pool their individual value orderings concerning the potential options in O , denoted by the profile $\mathcal{R} = \langle R_1, R_2, \dots, R_n \rangle$, in order to create or construct their collective, all-things-considered value ordering R . Similar to such an ordering for individuals, xRy denotes that group as a whole judges that option x is better than or equal to option y . P and I are then defined in terms of R in the same way that R_i and I_i were defined in terms of R_i .

The idea of pooling individual assessments into a single group assessment is generalized with the idea of a pooling function, which is *any* process taking a profile of value orderings and returning a single one. Arrow refers to these as 'social welfare functions'.¹⁶ I avoid this usage because it suggests that these functions are necessarily social and only for assessing welfare. Pooling functions certainly may apply in such situations, but their applicability is more general than that. Formally speaking, a pooling function is any function F that has the form $F(\mathcal{R}) = R$. So

¹⁶Arrow, *Social Choice and Individual Values*, p. 23.

while Arrow and others may use the language of 'social' choice, I emphasize again that this formal approach has farther reaching applications. For instance, it is relevant for the decision making of a single individual who has a pluralism of values and commitments, each of which provides a single ranking of the options, and who wishes to derive a single, all-things-considered judgment from them.¹⁷ For this chapter, however, I remain focused on the relevance of this apparatus for the liberal state. I do return to cases of individual decision making later (in section 4.3).

Now a non-trivial pooling function is one that involves more than one person (or more than one value ordering) and more than two potential options. Of course, in non-trivial cases, there are many possible ways to pool individual value orderings, especially since the number of ways increases dramatically as the number of people and potential options increases.¹⁸ In order to winnow possible pooling functions to a more manageable number, Arrow specifies six conditions that he claims are reasonable for a pooling function to satisfy if it is to be both rational and responsively impartial.¹⁹ The problem, however, is that Arrow's theorem demonstrates that no pooling function satisfies all of these conditions. While I present these conditions in this section with limited comment, I return to them in the following sections, fleshing them out in response to attempts to show that they are inappropriate or unnecessary in the context of decision making by the liberal state.

The first two conditions are the rationality postulates, which respectively require that the input and output of a pooling function be rational in accord with the default rationality thesis:

Individual Rationality ($\mathcal{IR}$): The pooling function only accepts profiles $\mathcal{R}$ where each individual value ordering R_i in $\mathcal{R}$ is a weak ordering of the potential options.²⁰

Social Rationality ($\mathcal{SR}$): The pooling function always returns an evaluation R that is itself a weak ordering of the potential options.

Together, $\mathcal{IR}$ and $\mathcal{SR}$ specify that a pooling function only works when given

¹⁷Surprisingly few commentators on social choice theory recognize this connection between individual and collective decision making. One notable exception, though, is Susan L. Hurley, *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989), pp. 226–231.

¹⁸Given just two individuals and two options, there are $4^9 = 262,144$ possible pooling functions taking a set of weak orderings and returning a single one. Increasing the number of options to three, means there are 14^{169} —more than a googol!—of potential pooling functions.

¹⁹For more on this interpretation of Arrow's strategy, see R. Duncan Luce and Howard Raiffa, *Games and Decisions: Introduction and Critical Survey* (New York: Wiley, 1957), pp. 328–331.

²⁰This follows from Axioms I and II in Arrow, *Social Choice and Individual Values*, p. 13.

rational input and, then, it will always return a rational, all-things considered judgment. With these two conditions, Arrow assumes that the decisions of a group must be held to the same standards of rationality as those applying to any individual person. Just as an individual is supposed to do, when the state is facing a choice concerning a set of actual options in A , it should choose from the set of options denoted by $C(A, R)$, which is understood in terms of optimization. I draw attention to this presumed symmetry between individual and collective decision making, for it is the subject of chapter 4.

Moving on, the remaining conditions specify a minimal sense of responsive impartiality for a pooling function. The first of these requires that a pooling function is responsive to the individual evaluations when they express an unanimous judgment concerning the relative worth of two options:

Weak Pareto Principle ($\mathcal{P}$): If for all R_i in $\mathcal{R}$, xP_iy holds for potential options x and y , then the pooling function must return a collective value ordering R satisfying xPy .²¹

Certainly when there is no conflict over a particular pair of options—everyone judges one option superior to another—the outcome of the pooling function should reflect this. $\mathcal{P}$ thereby ensures that the function preserves any unanimous evaluations of superiority that may appear in a profile of individual value orderings.

The next condition blocks the most partial of assessments by preventing a pooling function from always using one of the individual value orderings to determine the collective value ordering, effectively ignoring the judgments of the other individuals in the group:

Non-Dictatorship ($\mathcal{D}$): The pooling function cannot be dictatorial. That is, there is no individual d in N , such that for all profiles $\mathcal{R}$ and all potential options x and y , whenever xP_dy holds the pooling function always returns a collective value ordering R satisfying xPy .²²

If the pooling function is supposed to be responsive to all the inputted evaluations, there cannot be a single individual value ordering that determines the collective one, irrespective of all the other individual orderings. A pooling function violating

²¹Arrow, *Social Choice and Individual Values*, p. 96. For obvious reasons, this condition is also referred to as ‘unanimity’ in William H. Riker, *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice* (San Francisco: W. H. Freeman, 1982), p. 117 and Dennis C. Mueller, *Public Choice III* (Cambridge: Cambridge University Press, 2003), p. 583.

²²This follows from Definition 6 and Condition 5 in Arrow, *Social Choice and Individual Values*, p. 30.

this may hardly be considered to be pooling anything at all, since it is ultimately partial to only one person's value ordering.

The fifth condition recognizes that the individual value orderings may diverge and that the pooling function must be prepared to handle these situations:

Unrestricted Domain ($\mathcal{U}$): The domain of the pooling function is the set of all logically possible profiles $\mathcal{R}$.²³

This implies that a pooling function cannot preemptively restrict the contents of the profiles it will accept. Combined with $\mathcal{IR}$ and $\mathcal{SR}$, this means that a pooling must be prepared to accept any logical combinations of weak orderings forming a profile and to always return a weak ordering. That is, if the input to the process is rational, that is, any possible weak ordering of the options, then the output of it should be rational as well.

The sixth, and final, condition concerns consistency between profiles of individual value functions and the optimal options in $C(A, R)$:

Independence of Irrelevant Alternatives ($\mathcal{IIA}$): Consider profiles $\mathcal{R}$ and $\mathcal{R}'$. Let R and R' be the respective collective value orderings returned by the pooling function (i.e., $F(\mathcal{R}) = R$ and $F(\mathcal{R}') = R'$). When given an agenda of available options A , if both profiles agree with respect to this agenda (i.e., for each i , and all x and y in A , $xR_i y$ holds if and only if $xR'_i y$ holds), then $C(A, R) = C(A, R')$.²⁴

$\mathcal{IIA}$ entails that if the individual evaluations change concerning potential options in O that are not in A , while remaining the same for those options in A , then the collectively optimal options from A must also stay the same. In other words, how the pooling function selects optimal options from those 'on the agenda' cannot be influenced by irrelevant options, which are those options not 'on the agenda'.

Unfortunately, these six conditions effectively winnow down the possible non-trivial pooling functions to the null set:

Arrow's General Possibility Theorem: There is no non-trivial pooling function satisfying $\mathcal{IR}$, $\mathcal{SR}$, $\mathcal{P}$, $\mathcal{D}$, $\mathcal{U}$, and $\mathcal{IIA}$.²⁵

²³This is Condition 1' in *ibid.*, p. 96. It is also known as 'universal domain', for instance, in Amartya K. Sen, 'Rationality and Social Choice', *American Economic Review* 85, 1 (March 1995): 1–24, p. 4.

²⁴This is Condition 3 in Arrow, *Social Choice and Individual Values*, p. 27. Given its content, it is referred to as 'independence of infeasible alternatives' by Charles R. Plott, 'Axiomatic Social Choice Theory: An Overview and Interpretation', *American Journal of Political Science* 20, 3 (August 1976): 511–596, p. 521 and Peter C. Fishburn, *The Theory of Social Choice* (Princeton: Princeton University Press, 1973), pp. 6–8.

²⁵This is Theorem 2 in Arrow, *Social Choice and Individual Values*, p. 97. Three illuminating

Any method pooling more than one individual value orderings concerning more than two potential options must violate at least one of these conditions. In particular, this theorem suggests that there will always be instances where responsive impartiality leads to irrational decisions—even when the state’s constitutive members are themselves being reasonable. So the tension between responsive impartiality and state agency from section 1.4 reappears. Appealing to democratic procedures does not appear to resolve it after all. Before assessing this claim any further, I first present a second impossibility result which involves a slightly more sophisticated conception of rationality.

3.3 Bayesian Rationality

More recently, Teddy Seidenfeld, Joseph Kadane, and Mark Schervish together have shown that a similar conclusion is reached when rationality is understood in Bayesian terms.²⁶ This conception of Bayesian rationality is a refinement on the default rationality thesis. In this Bayesian analysis, the formation of an individual value ordering R_i is more nuanced, taking into account uncertainties in the consequences of carrying out or otherwise implementing an option. It then stipulates how to systematically use this information to generate a single, ‘Bayesian’ value function v_i for making all-things-considered judgments.

This analysis assumes there are p states of affairs, denoted by the set $\Omega = \{\omega_1, \omega_2, \dots, \omega_p\}$, each of which leads to certain consequences based on the option selected by the individual. The individual is uncertain as to which ω_j holds, but is certain what consequence occurs when an option is selected and a given state holds. Therefore, there are $m \times p$ consequences, where carrying out option k in state j leads to consequence $c_{k,j}$, which is represented by the decision matrix in figure 3.3. According to this framework, individual i has a pair of assessments, $(Pr_i, \bar{v}_i)$, where Pr_i is i ’s probability distribution over the states in Ω and $\bar{v}_i$ is i ’s von Neumann-Morgenstern value function providing cardinal assessments of the various consequences.²⁷ The Bayesian value function v_i , representing the ‘expected’

proofs may be found in Wulf Gaertner, *A Primer in Social Choice Theory* (Oxford: Oxford University Press, 2006), pp. 19–31.

²⁶Teddy Seidenfeld, Joseph B. Kadane and Mark J. Schervish, ‘On the Shared Preferences of Two Bayesian Decision Makers’, *Journal of Philosophy* 86, 5 (May 1989): 225–244.

²⁷For more on the specific properties of probability distributions and von Neumann-Morgenstern value functions in the Bayesian context, see, for instance, David M. Kreps, *Notes on the Theory of Choice* (Boulder, CO: Westview Press, 1988), pp. 43–52, 115–119, and Martin Peterson, *An Introduction to Decision Theory* (Cambridge: Cambridge University Press, 2009), pp. 94–103, 118–127.

	ω_1	ω_2	...	ω_j	...	ω_p
x_1	$c_{1,1}$	$c_{1,2}$		$c_{1,j}$		$c_{1,p}$
x_2	$c_{2,1}$	$c_{2,2}$		$c_{2,j}$		$c_{2,p}$
$\vdots$						
x_k	$c_{k,1}$	$c_{k,2}$		$c_{k,j}$		$c_{k,p}$
$\vdots$						
x_m	$c_{m,1}$	$c_{m,2}$		$c_{m,j}$		$c_{m,p}$

FIGURE 3.3: A decision matrix concerning m different potential options and p states of affairs with $m \times p$ consequences.

value of implementing each option, is then determined from these as follows:

$$v_i(x) = \sum_{j=1}^p (\Pr_i(\omega_j) \times \bar{v}_i(c_{x,j})), \text{ for each option } x.$$

$\Pr_i(\omega_j)$ is the probability that state ω_j holds, and $\bar{v}_i(c_{x,j})$ is the cardinal valuation of the consequences for implementing option x when state ω_j holds. Finally, a value ordering R_i is generated: xR_iy if and only if x does not yield less expected value than y (that is, $v_i(x) \geq v_i(y)$). The individual then should select the optimal option according to R_i , and thereby maximize expected value. This is the essence of Bayesian rationality.

Despite its seemingly complexities, this account of rationality is a relatively simple extension of Kenneth Arrow's, allowing for uncertainty and handles it in a Bayesian fashion. Given the prevalence of uncertainties in the world, this understanding of rationality seems plausible for making decisions under uncertainty. Furthermore, similar to the default rationality thesis, this Bayesian account does not constrain what values and commitments a person may employ to evaluate the consequences. It just specifies how these assessments and probability judgments should come together to form an all-things-considered judgment of the options.

In addition, this Bayesian notion of rationality may apply to state decisions as well. Again, I draw attention to the idea that the decisions of both the individual and the liberal state ought to be judged by the same norms of rationality. In this case, those norms now reflect augmenting the default rationality thesis in Bayesian directions. So as with individuals, a weak ordering R is generated from a single probability/value pair $(\Pr, \bar{v})$ that represents society's overall probability distribution over states and utility over outcomes. The state then uses this ordering to select an option by optimization.

Once again, if the state seeks to pool the judgments of its constitutive members to determine its overall assessment of the options, some mechanism may be

employed to pool a collection of probability/value pairs, denoted by the profile $\mathcal{PV} = \langle (\Pr_1, \bar{v}_1), (\Pr_2, \bar{v}_2), \dots, (\Pr_n, \bar{v}_n) \rangle$, into a single probability/value pair $(\Pr, \bar{v})$.²⁸ As Arrow does with pooling functions, I generalize this by defining a Bayesian pooling function as a mechanism that accepts some $\mathcal{PV}$ and returns a single $(\Pr, \bar{v})$ for determining which option the state should select.

Seidenfeld et al. propose certain conditions for Bayesian pooling functions to satisfy, which are similar to Arrow's. These conditions are as follows:

Individual Rationality ($\mathcal{IR}_B$): The Bayesian pooling function only accepts profiles $\mathcal{PV}$ where each individual probability/value pair $(\Pr_i, \bar{v}_i)$ generates a weak ordering of potential options.

Social Rationality ($\mathcal{SR}_B$): The Bayesian pooling function must always return a single probability/value pair $(\Pr, \bar{v})$ that itself generates a weak ordering of the potential options.

Weak Pareto Principle ($\mathcal{P}_B$): If for all $(\Pr_i, \bar{v}_i)$ in $\mathcal{PV}$, $(\Pr_i, \bar{v}_i)$ entails that xP_iy , then the Bayesian pooling function must return a collective probability/value pair $(\Pr, \bar{v})$ that entails xPy .

Non-Dictatorship ($\mathcal{D}_B$): The Bayesian pooling function cannot be dictatorial. That is, there is no individual d in N , such that for all profiles $\mathcal{PV}$, $(\Pr_d, \bar{v}_d)$ is equivalent to $(\Pr, \bar{v})$.

Unrestricted Domain ($\mathcal{U}_B$): The domain of the Bayesian pooling function is the set of all logically possible profiles $\mathcal{PV}$.

Seidenfeld et al. then prove the following:

Seidenfeld-Kadane-Schervish Theorem: There is no non-trivial Bayesian pooling function that satisfies $\mathcal{IR}_B$, $\mathcal{SR}_B$, $\mathcal{P}_B$, $\mathcal{D}_B$, and $\mathcal{U}_B$.²⁹

Therefore, even when accounting for uncertainties in reasoning, it appears that the liberal state is unable to always choose both rationally and in accord with

²⁸Note that this definition of $\mathcal{PV}$ entails that each individual may have a different probability distribution over the states, as well as different value assessments concerning the consequences, perhaps owing to different levels of expertise in relevant areas that might aid in predicting these states.

²⁹Part (i) of Theorem 1 in Seidenfeld, Kadane and Schervish, 'Shared Preferences of Two Decision Makers', p. 236. Strictly speaking, this theorem only shows the impossibility when there are exactly two individuals involved. However, this result has been generalized by Jay Goodman, 'Existence of Compromises in Simple Group Decisions', PhD thesis (Pittsburgh, PA: Department of Statistics, Carnegie Mellon University, 1988), and Philippe Mongin, 'Consistent Bayesian Aggregation', *Journal of Economic Theory* 66, 2 (August 1995): 313–351.

responsive impartiality. In fact, this result appears somewhat more troubling than Arrow's insofar as there is no assumption of independence of irrelevant alternatives (IIA).

Though having now laid out these two impossibility theorems, I focus primarily on Arrow's since it has generated by far more commentary. Regardless, I now move on to discuss many of the proposed ways to avoid the pessimistic implications for the liberal state these theorems pose. While I find most of them wanting, I believe that a more radical break with the underlying rationality assumptions shared by both impossibility theorems is necessary for liberal democracy to truly avoid their implications. The idea in the following sections is therefore to show why such a move is necessary.

3.4 The Rationality Postulates

A pooling function satisfying the rationality postulates—individual rationality ($\mathcal{IR}$) and social rationality ($\mathcal{SR}$)—is simply a mechanism for constructing a single weak ordering from a set of them. While these conditions may be in accord with the default rationality thesis, it might still be perplexing at first to require a pooling function for collective decision making to satisfy them. At first glance, it certainly seems that all that is needed for decision making in these cases is designating some option(s) as the 'best' option(s). Returning the full weak ordering is apparently unnecessary and therefore so is $\mathcal{SR}$. Furthermore, common political mechanisms for pooling individual judgments such as the majority and plurality voting rules often do not solicit any more than the option each person judges as the best, and thus they violate $\mathcal{IR}$.

Rejecting $\mathcal{SR}$, though, sparks a theoretical problem. Consider the profile of individual value orderings illustrated in figure 3.4. Say the group employs a pooling method that violates $\mathcal{SR}$ insofar it selects coal power as the winner, but gives no further information concerning the other options. If coal power is for whatever reason suddenly no longer an option it is then unknown what power plant the group judges as second best. A commitment to choosing the best option given the individuals' judgments seems to imply an implicit commitment to an entire rank ordering of the alternatives; should the best become unavailable, a pooling system must choose the second best, and so on. This disposition suggests that pooling functions should return an entire collective value ordering satisfying $\mathcal{SR}$. A similar concern also motivates $\mathcal{IR}$. If a pooling method only solicits each person's most favored option, the removal of coal power makes the judgments from John and Paul worthless. Another solicitation is then necessary to determine the option they now rank first. Obtaining full individual value orderings, as dictated by $\mathcal{IR}$,

	John	Paul	George	Stuart	Pete
1 st	coal	coal	nuclear	hydro	solar
2 nd	nuclear	nuclear	solar	nuclear	nuclear
3 rd	hydro	hydro	hydro	solar	hydro
4 th	solar	solar	coal	coal	coal

FIGURE 3.4: *A group's profile of individual value orderings for choosing from coal power, nuclear power, hydropower, and solar power.*

simply solves this problem.

Of course, when there are only two options, this is not a real issue for majority rule, which selects the option with over half of the first-place votes. $\mathcal{IR}$ and $\mathcal{SR}$ are implicitly satisfied: once it is known what option is best according to an ordering—regardless of whether it is an individual or the collective one—the worst option according to that ordering is also revealed. But an immediate problem arises when there are more than two options because it is possible that there is no majority placing the same option on top of their evaluations.

In response, the plurality method may be employed, where the option with the most first-place votes wins. But this method faces its own problems because of its rejection of $\mathcal{IR}$. For instance, there is no majority winner in figure 3.4, whereas the plurality method selects coal power as the best option. The concern is that since the plurality method does not use all the information contained the individual value orderings, it ignores the fact that an overwhelming majority of people judge both nuclear power and hydropower as superior to coal power.³⁰ Pointing out this concern, Jean-Charles Borda claims the plurality method is not very appealing:

The voters cannot give a sufficiently complete account of their opinions of the candidates. . . . If a form of election is to be just, the voters must be able to rank each candidate according to his merits, compared successively to the merits of each of the others.³¹

According to Borda, $\mathcal{IR}$ ensures that a pooling method has such ‘sufficiently complete accounts’.

Borda along with his compatriot and often interlocutor, the Marquis de Condorcet, nevertheless struggle to devise pooling methods satisfying the rationality

³⁰Margaret Thatcher's Conservative party in the United Kingdom might have been aware of this fact in calling for an election in 1983 before the opposition—the Labour party and the alliance of Liberals and Social Democrats—could organize a unified front. At that time, a majority of voters favored more liberal candidates than conservative ones, but since these votes were scattered across different candidates, the plurality method brought more victories to the Conservative party.

³¹*On Elections by Ballot* from McLean and Urken, eds., trans., *Classics of Social Choice*, p. 84.

postulates. Condorcet proposes a rule that orders each pair of options based on how a majority of individuals ranks that pair, assembling a collective evaluation from these pairwise comparisons:

The Condorcet Rule: xRy if and only if the number of individual evaluations R_i in $\mathcal{R}$ with xR_iy is at least as great as the number of those with yR_ix .

When used for as a method for pooling value orderings, it is impossible for there to exist a top ranked ‘Condorcet winner’ that a majority of individuals judges as inferior to some other option. This stands as an improvement to the plurality method. For in the case of the profile from figure 3.4, the Condorcet rule ranks nuclear power highest with hydropower second, solar power third, and coal power last.³² In fact, this rule satisfies $\mathcal{IR}$, weak Pareto ($\mathcal{P}$), non-dictatorship ($\mathcal{D}$), unrestricted domain ($\mathcal{U}$), and independence of irrelevant alternatives ($\mathcal{IIA}$). Consequently, as entailed by Kenneth Arrow’s theorem, it must violate $\mathcal{SR}$.

Condorcet was aware of this failing with his method but was unable to solve it—unsurprising to us given Arrow’s theorem, but frustrating to him given his ignorance of the result. Consider the profile of individual evaluations in figure 3.5. In this situation a majority (composed of Dewey and Cheetham) ranks coal power over nuclear power; another majority (Dewey and Howe) ranks nuclear power over hydropower; and yet another majority (Cheetham and Howe) ranks hydropower over coal power. This cyclic ranking of options is known as Condorcet’s paradox. No matter which option is selected, there is a majority preferring another option to it. Therefore, there is no top-ranked, optimal, winning alternative for the group to choose.³³ Furthermore, as seen already, no weak ordering can represent cyclic comparisons. Under circumstances like these, the Condorcet rule cannot return an evaluation satisfying $\mathcal{SR}$.

Meanwhile, Borda offers his own ‘system of marks’ that ranks an option based on how many pairwise comparisons it wins across all the individual evaluations against all the other options:

³²A majority (George, Stuart, and Pete) ranks nuclear power, hydropower, and solar power all over coal power; another majority (John, Paul, and George) ranks nuclear power over both hydropower and solar power; and a third majority (John, Paul, and Stuart) ranks hydropower over solar power.

³³This point is often neglected. William Riker, whose formal rigor is otherwise admirable, suggests that a cycle implies that society is indifferent between these options, William H. Riker, *Democracy in the United States*, 2nd ed. (London: Macmillan, 1964), p. 120. Since indifference is an acceptable way for a pooling function to order options, this implies that the Condorcet rule *does* return an ordering. However, this is not the case. In figure 3.5, according to the Condorcet rule, the group is *not* indifferent between coal and hydropower: coal power is judged *superior* to hydropower (by transitivity) and vice-versa (by a pairwise comparison). Neither judgment is one of indifference, and both together entail a violation of the asymmetry of superiority.

	Dewey	Cheetham	Howe
1 st	coal	hydro	nuclear
2 nd	nuclear	coal	hydro
3 rd	hydro	nuclear	coal

FIGURE 3.5: Dewey, Cheetham, and Howe's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower.

The Narrow Borda Rule: Suppose there are a k number of available options on the agenda A . For each top position (only relative to the other options on the agenda) held in an R_i , an option receives $k - 1$ points; for each second place position held (only relative to those on the agenda), it gets $k - 2$ points; $\dots$; for each bottom position held (only relative to those on the agenda), it gets 0 points. xRy if and only if x does not score lower than y .³⁴

So whereas the Condorcet rule ranks two options relative to each other solely based on the pairwise comparisons between those two options, the narrow Borda rule rank two options relative to each other based on pairwise comparisons between those two options as well as how those two options compare to *all* the other options.³⁵

To illustrate, consider the situation in figure 3.6. In this case, the Condorcet rule ranks coal power over nuclear power based on the fact that more people judge coal power superior to nuclear power than the other way around.³⁶ The presence of

³⁴I present another version of the Borda rule in section 3.6 (also see n. 55 in this chapter).

³⁵When comparing the Condorcet and narrow Borda rules, commentators sometimes misleadingly call Condorcet a pairwise method and Borda a relational one. Both rules, however, are pairwise. To illustrate this, first note that the narrow Borda rule is almost identical to Sister Natana's method from *Blanquerna*. As is, the latter method violates $\mathcal{IR}$, which is illustrated by the following individual ballot (based on a larger ballot Ramon Lull presents in *The Art of Elections* (1299), translated in McLean and Urken, eds., trans., *Classics of Social Choice*, p. 74):

	Natana	Maria
Sarah	<i>Sarah</i>	<i>Maria</i>
Maria	<i>Natana</i>	

In this case there are three candidates, Sister Natana, Sister Maria, and Sister Sarah, and so this ballot requires three pairwise comparisons. The rows and columns denote the candidates and each cell denotes whom the voter favors in a pairwise comparison between the row and column candidates. This ballot has been filled out irrationally because it denotes an irrational cycle: Sister Natana is superior to Sister Maria, who is superior to Sister Sarah, who is superior to Sister Natana. Amending this method to require that individuals submit ballots reflecting a weak ordering of the options, then makes it identical to the narrow Borda rule. In that case, the number of pairwise comparisons a candidate wins on such ballots is equal to the number of points that a candidate would receive according to the narrow Borda rule.

³⁶A majority, made up of Dewey and Howe, places coal power over nuclear power; another

	Dewey	Cheetham	Howe	Condorcet	Narrow Borda
1 st	coal	nuclear	hydro	coal	coal
2 nd	nuclear	coal	coal	nuclear	nuclear
3 rd	hydro	hydro	nuclear	hydro	hydro

FIGURE 3.6: Dewey, Cheetham, and Howe's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower. Using the Condorcet and narrow Borda rules on this profile, with the agenda A the same as the set of options O , yields identical value orderings.

hydropower in the individual value orderings does not play a role for this particular judgment. Given the same profile and all three options on the agenda, the narrow Borda rule ranks coal power over nuclear power based on the total number of pairwise comparisons each option wins, including those comparisons involving hydropower. For instance, coal power wins a total of four comparisons, which is seen in the three value ordering:

- (1) With Dewey's ordering, coal power wins two comparisons, one when compared with nuclear power and one when compared with hydropower,
- (2) With Cheetham's ordering, coal power wins one additional comparison because it wins one when compared with hydropower and none when compared with nuclear power, and finally,
- (3) With Howe's ordering, coal power wins one more comparison, since it wins one versus nuclear power and none when compared with hydropower.

Computing according to the formal version of the narrow Borda rule confirms that coal power earns a total of four points. Similar calculations show that nuclear power scores a total of three points. Notice that unlike the Condorcet rule, the narrow Borda rule is sensitive to the placement of hydropower in the individual orderings. As the above calculations show, the placement of hydropower plays a critical role in determining the number of points of the other two options receive. Even so, the narrow Borda rule ultimately orders the options the same as the Condorcet rule: nuclear power is the 'narrow Borda winner', with hydropower second and coal power third.³⁷

The narrow Borda rule does not always agree with the Condorcet rule, however. Consider the situation in figure 3.7. In that case, the Condorcet rule ranks hydro-

majority, made up of Dewey and Cheetham, places coal power over hydropower and nuclear power over hydropower.

³⁷Coal power scores $2 + 1 + 1 = 4$ points, nuclear power scores $1 + 2 + 0 = 3$ points, and hydropower scores $0 + 0 + 2 = 2$ points.

	John	Paul	George	Stuart	Pete
1 st	hydro	hydro	hydro	coal	coal
2 nd	coal	coal	coal	nuclear	nuclear
3 rd	nuclear	nuclear	nuclear	hydro	hydro

	Condorcet	Narrow Borda
1 st	hydro	coal
2 nd	coal	hydro
3 rd	nuclear	nuclear

FIGURE 3.7: *A group's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower. Using the Condorcet and narrow Borda rules on this profile, with the agenda A the same as the set of options O , yields conflicting value orderings.*

power over coal power while the narrow Borda rule ranks them the other way.³⁸ Divergence between these two methods occurs here because the Condorcet winner is ranked low by several individual evaluations, causing the narrow Borda rule to rank it lower. So a majority can sometimes overturn a narrow Borda winner—in this case, coal power in favor of hydropower. On the positive side, though, the narrow Borda rule succeeds where the Condorcet rule fails: it returns an ordering satisfying $\mathcal{SR}$ for the profile in figure 3.5 when all three options are on the agenda, with all three options tying.³⁹

Despite this, the narrow Borda rule has its own difficulties. Consider the profile of individual evaluations in figure 3.8. When all four options are in agenda A , the narrow Borda rule ranks nuclear power over coal power. But when solar power is taken off the agenda, as in A' , the narrow Borda rule suddenly ranks coal power over nuclear power.⁴⁰ This should not be too surprising: we have already seen that the points an option earns depends critically on how it fares relative to all the other options on that particular agenda. This is a violation of $\mathcal{SR}$; in particular, the asymmetry of superiority is violated, causing there to be no single weak ordering that can satisfy the two comparisons given by the narrow Borda rule as the agenda changes. It turns out that, like the Condorcet rule, the narrow

³⁸The Condorcet rule's ordering comes exclusively from the majority coalition formed between John, Paul, and George. When all three options on the agenda, then according to the narrow Borda rule hydropower scores $2 + 2 + 2 + 0 + 0 = 6$ points, coal power scores $1 + 1 + 1 + 2 + 2 = 7$ points, and nuclear power scores $0 + 0 + 0 + 1 + 1 = 2$ points.

³⁹Each option earns 6 points apiece.

⁴⁰With all four options on the agenda, nuclear power scores $2 + 3 + 0 + 0 + 3 = 8$ points while coal power scores $3 + 1 + 1 + 2 + 0 = 7$ points. With solar power off the agenda, coal power now scores $2 + 1 + 1 + 2 + 0 = 6$ points and nuclear power scores $1 + 2 + 0 + 0 + 2 = 5$ points.

	Dewey	Cheetham	Howe	Argue	Phibbs
1 st	coal	nuclear	solar	solar	nuclear
2 nd	nuclear	solar	hydro	coal	solar
3 rd	solar	coal	coal	hydro	hydro
4 th	hydro	hydro	nuclear	nuclear	coal

Narrow Borda on Agenda . . .		
	A	A'
1 st	solar	coal
2 nd	nuclear	nuclear
3 rd	coal	hydro
4 th	hydro	—

FIGURE 3.8: *A group's profile of individual value orderings for choosing from coal power, nuclear power, hydropower, and solar power. If agenda $A = \{\text{coal, nuclear, hydro, solar}\}$ and $A' = \{\text{coal, nuclear, hydro}\}$, then using the narrow Borda rule on this profile on these different agendas leads to the value orderings shown above.*

Borda rule satisfies $\mathcal{IR}$, $\mathcal{P}$, $\mathcal{D}$, $\mathcal{U}$, and $\mathcal{JJA}$, but, as this case illustrates, violates $\mathcal{SR}$.

Surprisingly, most analyses of the impossibility results of social choice theory do not challenge the rationality postulates. The default rationality thesis seems to have a firm hold, and so most critics try to find ways to ensure that rules like Condorcet and narrow Borda construct a weak ordering over the options. This is a shame, for in the next chapter, I will return to the default rationality thesis, arguing that it should be overturned. Rejecting this thesis then undermines the heart of the impossibility results. At this point, however, such a move might appear too extreme, so I must first show why other, more common, moves against the theorems are not satisfactory.

3.5 The Condorcet Rule and Domain Restriction

While the Condorcet rule suffers from the inability to always return a collective value ordering satisfying social rationality ($\mathcal{SR}$), it turns out that there exist ways around this problem. Such ways involve restricting the set of permissible profiles of individual value orderings in order to exclude those ‘problem’ profiles that confound the Condorcet rule, causing this rule to make cyclical judgments. Doing so does come at the expense of unrestricted domain ($\mathcal{U}$), which requires that a pooling function handle all logically possible profiles. On the face of it, however, $\mathcal{U}$ does appear an attractive condition for the liberal state because it reinforces the ideal of respecting each individual’s liberty to freely order the options—no value

orderings are imposed on the state's constitutive members. The fact of reasonable pluralism allows different individuals to evaluate the options differently, and so it seems the liberal state should then be prepared to handle any profile of value orderings and return a collective one.

Nevertheless, there are 'structured' profiles that always ensure that the Condorcet rule always return a collective ordering satisfying $\mathcal{SR}$. If the domain of profiles may permissibility be restricted to a sets of these, then the dire implications of Kenneth Arrow's theorem for the Condorcet rule are avoided.⁴¹

One simple type of structured profile occurs when everyone in the group reaches a consensus on the rankings of the options. When advocating an approach to democracy based upon 'public debate and confrontation', John Elster rather ambitiously suggests that this approach to decision making by the state would generate such concordant profiles because 'a rational discussion would tend to produce unanimous preferences'.⁴² Were such substantive agreements possible, there would be no need for any pooling scheme at all. The collective judgment is simply identical to that shared by each individual.⁴³ Unfortunately, there is little reason and no empirical evidence to believe that rational discussions have this desired effect. For instance, even at its very best, the United States Supreme Court rarely reaches unanimous consensus on judgments—and this is a rather small body of decision makers. At the societal level, the fact of reasonable pluralism accepts that this possibility might be even more unlikely. As a result, few liberal democrats currently support Elster's position.

Another paradigmatic type of structured profile is Duncan Black's notion of single-peaked profiles. A profile of orderings is single-peaked whenever the following circumstances obtain:

- (1) There exists a left-to-right axis ordering the options, where
- (2) Each individual evaluation has a most highly valued point (that is, the option it ranks first) on this axis, and

⁴¹Unfortunately, trying to similarly preclude the 'problem' profiles for the narrow Borda rule is not so easy, as most profiles are problematic for it. This is because, for virtually any profile, the narrow Borda rule runs into problems as the available options on the agenda A change.

⁴²Jon Elster, 'The Market and the Forum: Three Varieties of Political Theory', in *The Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 103–132, p. 112. Elster's rationale is that democratic deliberation would tend to produce 'informed and other regarding preferences'. Elster, however, fails to justify why such judgments must then be unanimous. Certainly two extreme altruists could disagree over how to allocate resources: each arguing that the other should get the most.

⁴³As Arrow, *Social Choice and Individual Values*, p. 74, puts it, in the case of complete unanimity, a pooling function could just pick any individual and use his or her evaluation as the collective one. This, of course, violates non-dictatorship ($\mathcal{D}$); but with the assumption of unanimity, $\mathcal{D}$ has little normative force.

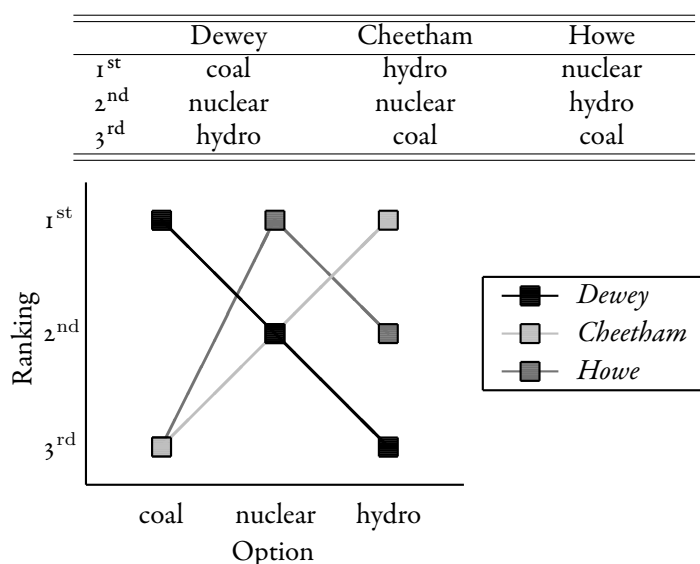


FIGURE 3.9: Dewey, Cheetham, and Howe's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower. This profile is also single-peaked, which is seen in the accompanying graph, which displays a left-to-right continuum based on the expenses of each power plant, where coal power is the cheapest power source, hydropower is the most expensive, and nuclear power is in the middle.

- (3) Each evaluation only decreases as options are further from that point.

Consider the profile of individual value orderings in figure 3.9, which is rendered on an axis ordering the power plants from left to right in terms of increasing overall financial impact on the community. Here we see that coal power is Dewey's most highly valued point because he judges it best to spend the least amount of money on a power plant. As actions move further from coal power on the expense axis, Dewey's rankings for the other options decrease accordingly. Cheetham's and Howe's respective judgments also have a single most valued point on this axis, and therefore this entire profile is single-peaked.

Black proves that single-peaked profiles possess an interesting property:

Black's Possibility Theorem for Single-Peaked Profiles: Given an odd number of individuals in society N and single-peaked profiles of individual value orderings, the Condorcet rule satisfies $\mathcal{IR}$, $\mathcal{SR}$, $\mathcal{P}$, $\mathcal{D}$, and $\mathcal{JJA}$.⁴⁴

⁴⁴Duncan Black, *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958), pp. 16–18; and Arrow, *Social Choice and Individual Values*, p. 78. The demand for an odd number of people is actually trivial. If there is an even number of options under a single-peaked profile, it is possible for there to be multiple optimal options, Riker, *Liberalism Against Populism*,

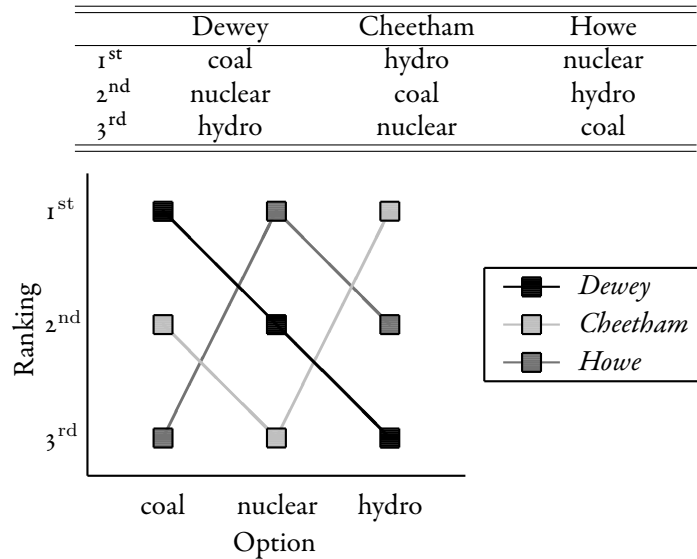


FIGURE 3.10: Dewey, Cheetham, and Howe's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower. This is the same as in figure 3.9, except that Cheetham's value ordering is different. The result is that this profile is *not* single-peaked, which is demonstrated in the accompanying graph, which displays the same continuum based on expenses as done in figure 3.9.

If its domain is restricted to single-peaked profiles, then the Condorcet rule satisfies all of Arrow's conditions but $\mathcal{U}$. For example, if the Condorcet rule is employed for the profile in figure 3.9, nuclear power is the definitive Condorcet winner. On the other hand, applying the Condorcet rule to a multi-peaked profile may provide no optimal alternative, as is the case for Condorcet's paradox seen in figure 3.10.⁴⁵

The viability of this result then turns on whether it is permissible for individuals' evaluations to be so restricted. David Miller, John Dryzek, and Christian List argue that public deliberation implicitly induces deliberators to reach a single-peaked profile.⁴⁶ Deliberators are supposed to appeal to interests general to all the other participants in order to justify their particular value orderings to each other.

p. 126. This is harmless because no cycles will emerge and indifference is acceptable according to the rationality postulates.

⁴⁵Black's theorem does not prove that *all* multi-peaked profiles cause the Condorcet rule to fail to return a collective evaluation satisfying $\mathcal{SR}$. Rather, the Condorcet rule can handle some multi-peaked profiles.

⁴⁶David Miller, 'Deliberative Democracy and Social Choice', *Political Studies* 40, special issue (August 1992): 54–67; John S. Dryzek, *Deliberative Democracy and Beyond: Liberals, Critics, Contestations* (Oxford: Oxford University Press, 2000); and John S. Dryzek and Christian List, 'Social Choice Theory and Deliberative Democracy: A Reconciliation', *British Journal of Political Science* 33, 1 (January 2003): 1–28, 'Social Choice Theory and Deliberative Democracy: A Response to

Dryzek and List suggest that this process should lead deliberators to understand the issue in terms of a single generalizable interest. Miller suspects that this occurs because deliberation activates norms that the participants should already implicitly share, or it allows the participants to create new norms for that situation. With a single perspective or interest identified, Dryzek and List claim that deliberation should elicit agreement among participants concerning a single axis representing this interest and how the alternatives are aligned from left to right on this axis. Finally, Dryzek and List insinuate, without argument, that rationality should direct deliberators toward a single-peaked profile on this axis: 'it might be argued that rationality requires one determinate dimension on which preferences are single-peaked'.⁴⁷ Employing the Condorcet Rule following public deliberation apparently ensures a result satisfying $\mathcal{SR}$.

While this is certainly an improvement over John Elster's claims about deliberation, I find it quite implausible that rationality ought to demand what Dryzek and List claim. Certainly the default rationality thesis does not demand it. In some, and perhaps even most, circumstances there is nothing irrational in adopting an evaluation that is multi-peaked on some recognized axis—even when only that axis is at stake. For instance, Cheetham's multi-peaked ranking in figure 3.10 could be perfectly reasonable; Cheetham maintaining that if a power plant is to be built, best not be stingy with the process, second-best to spend as little as possible. This type of position is not limited to axes involving costs either. It could reasonably occur in many different circumstances, summed up by the position 'best to do it right or not at all' or 'in for a penny, in for a pound'.⁴⁸ This is hardly irrational, and therefore it is hardly justifiable on grounds of rationality to cause individuals to adopt single-peaked orderings.

Furthermore, given the complexity of most policy issues, it is dubious that deliberators would, or even should, consistently agree on a single axis to judge the issue. In contemporary political debates, there is no shortage of instances where either several irreducible interests are at stake or a single interest does not reduce to a single axis. Nevertheless, Dryzek and List presume that deliberators will often

Aldred', *British Journal of Political Science* 34, 4 (October 2004): 752–758.

⁴⁷'Social Choice Theory and Deliberative Democracy: A Reconciliation', p. 15. I also find no argument for this position in Riker, *Liberalism Against Populism*, and Mueller, *Public Choice III*, whom Dryzek and List cite as agreeing with this position.

⁴⁸Ross Harrison, *Democracy* (London: Routledge, 1993), pp. 203–204, advances this type of reasoning. A historical example of such reasoning might have occurred during the Vietnam War, where many believed it was quite reasonable to prefer full-scale war to no involvement to moderate involvement. For another, though perhaps less convincing, argument against the claim that rationality requires single-peaked value orderings, see Jonathan Aldred, 'Social Choice Theory and Deliberative Democracy: A Comment', *British Journal of Political Science* 34, 4 (October 2004): 747–752.

reach agreement on these contentious issues. They provide anecdotes and speculation, but nothing conclusive for why deliberation ensures these agreements.⁴⁹ Nor do they prescribe concrete proposals for structuring deliberation to do so. Without these, it is difficult to believe that if participants disagree in their overall evaluations of the options, they would naturally agree on these other issues.

One could argue that in fact people always, or almost always, really have single peaked valuations but misconceive their own values, and that structured conversation—or institutions of democratic deliberation—brings individuals to recognize and explicitly endorse their own true ordering. There is, to the best of my knowledge, no evidence for such inner harmony, and I have no idea how any could be obtained.

Apart from single-peaked profiles, profiles may be restricted in other ways that also ensure that the Condorcet rule satisfies SR. Two prominent examples are Richard Niemi's 'sufficiently single-peaked' profiles or Amartya Sen's 'value-restricted' profiles.⁵⁰ As with single-peaked profiles, there is little reason to believe that deliberation produces these either, or that it would be good if it did. The hope that people will naturally restrict the form of their value orderings is too fragile a possibility upon which to rest the success of making collective judgments that are

⁴⁹In defending their account, Dryzek and List appeal to empirical evidence from Christian List, Iain S. McLean, James S. Fishkin and Robert C. Luskin, 'Deliberation, Preference Structuration, and Cycles: Evidence from Deliberative Polls' (unpublished manuscript, 2001); and Christian List, Robert C. Luskin, James S. Fishkin and Iain S. McLean, 'Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls' (unpublished manuscript, 2006); that appears to demonstrate that, in practice, deliberative opinion polls induce more single-peak profiles. As conceived by James S. Fishkin, deliberative opinion polls capture 'what the public would think, if it had a more adequate chance to think about the questions at issue', James S. Fishkin, *Democracy and Deliberation: New Directions for Democratic Reform* (New Haven, CT: Yale University Press, 1991), p. 1. That is, a deliberative poll takes a random sample of the population and brings them together to actually deliberate the issues before polling them on their opinions. This evidence is currently very limited, however, and it does little to substantiate a deliberative account of how single-peak profiles result from deliberation. That is, there is no ascertaining from the data what meaningful axis, if any, the participants used to evaluate the issue and organize their valuations. List (in personal correspondence) also concedes that no cyclic profiles were present before the deliberations they tracked, and so their evidence cannot show whether deliberation actually 'corrects' cyclic profiles. All it does show is an increase in 'proximity to single-peakedness' after deliberation, but this does not remove the concern that this did not result from more problematic sources such as groupthink, an aggressive demagogue dominating deliberation, overly passive participants, or unintentional, yet manipulative, issue framing by the polls' moderators.

⁵⁰Richard G. Niemi, 'Majority Decision-Making with Partial Unidimensionality', *American Political Science Review* 63, 2 (June 1969): 488–497; and Amartya K. Sen, 'A Possibility Theorem on Majority Decisions', *Econometrica* 34, 2 (April 1966): 491–499. Both contain possibility theorems similar to that for single-peaked profiles, but applying to sufficiently single-peaked and value-restricted profiles respectively.

not cyclical.

If people do not restrict their orderings voluntarily, it is always possible to either coerce them to do so or merely exclude their orderings from the process. Certainly it is neither unusual nor problematic to coerce or exclude those with particularly pernicious valuations; a person who votes for the enslaving his neighbors is hardly someone whose value ordering we should solicit (especially if you are his neighbor). But these are substantive restrictions, not structural ones, and even excluding uncontroversially immoral orderings is no guarantee the resultant profile will be single-peaked, sufficiently single-peaked, value-restricted, or of some other form that ensures a rational social evaluation. There is no need to settle whether coercion may sometimes trump libertarian values to conclude that the mere goal of being able to construct a collective ordering in accord with $\mathcal{SR}$ is of itself insufficient reason for coercion or rejection of $\mathcal{U}$.

3.5.1 Multidimensional Structuration

Suggesting that a complex issue be reduced to a one-dimension problem may demand that participants oversimplify the problem they are facing by focusing on a single axis, despite the fact that many other axes may be equally salient. Even in the simplified case of building a power plant, one could imagine many relevant axes to consider: building costs, services costs, environmental impact, economic impact, and so on. When focusing on a single axis, individuals could make decisions that are, all-things- (or all-axes-) considered, not normatively appealing.

John Dryzek and Christian List, however, do recognize that given the complexity of many policy issues, it is dubious that rationality would always induce a single-peaked profile to emerge on an issue, and thus prevent the occurrence of cyclic collective judgments under the Condorcet rule. Cases exist where either several irreducible generalizable interests are at stake or a single generalizable interest cannot be reduced to a single dimension. In situations like this, Dryzek and List suggest that it is these multiple incommensurable dimensions underlying an issue that are responsible for causing the Condorcet rule to report cyclic collective judgments. Nevertheless, they contend that deliberation still successfully mitigates these situations by restricting the domain through a process of what they call 'multidimensional preference structuration'.⁵¹

Dryzek and List claim that deliberation leads participants to identify the relevant dimensions for a given issue. Following that, each person may then 'disaggregate' the issue along these dimensions. That is, if there are k dimensions

⁵¹Dryzek and List, 'Social Choice Theory and Deliberative Democracy: A Reconciliation', pp. 18–21.

concerning the issue, disaggregation has the person identify how the following two sources of information come together to form her individual value ordering:

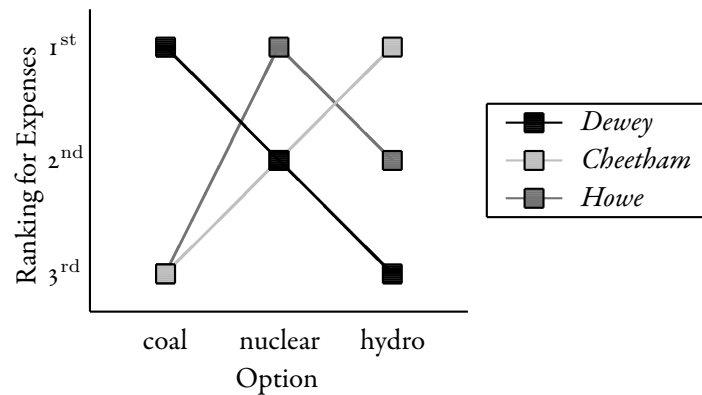
- (1) A profile of k value orderings, where each ordering is the individual's ranking of the options under a dimension, and
- (2) A set of k weights, where each weight represents the relative importance that individual attaches to a dimension.

Finally, Dryzek and List assert that deliberation should lead to what they call 'intra-dimensional single-peakedness': the profiles of individual value orderings *for each of these k dimensions* should all be single-peaked. As a result, the Condorcet rule now successfully generates a collective ordering *for each of these dimensions* satisfying the social rationality (SR) condition. With these orderings in place, Dryzek and List believe that deliberators may then reach an agreement on the weighing of the dimensions for making a final decision, or engage in logrolling to form a hierarchy of dimensions by allowing participants to indicate what weight they would accept on other dimensions in return for acquiescence on their preferred dimensions.

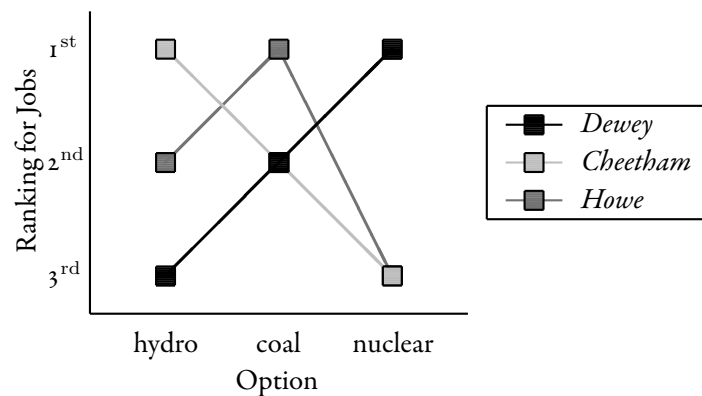
To illustrate this process of multidimensional preference structuration, recall the multi-peaked profile from figure 3.10. Applying the Condorcet rule directly to this profile yields cyclic judgments. According to Dryzek and List, deliberation might identify, for instance, two relevant dimensions for this issue: expenses and job creation. Further deliberation might then induce single-peaked profiles for each of these dimensions, as illustrated in figure 3.11(A) and (B). If Dewey and Howe give most weight to the dimension concerning expenses, while Cheetham assigns most weight to job creation, the final multi-peaked profile from figure 3.10 is not so mysterious. Since the individual dimensions are single-peaked, though, the Condorcet rule then returns the collective orderings presented in figure 3.11(C). Now all that remains is for the group to negotiate over how to weigh these two dimensions for making a final decision.

Unfortunately, it is this last part that Dryzek and List breeze over far too quickly. Once participants have arrived at the k value orderings representing each dimension, a significant problem reemerges. Recall that I emphasized (in section 3.2) the general nature of Arrow's theorem. In particular, it applies for any attempt to pool a collection of orderings into a single one. So just as the theorem applies when n individuals are attempting to pool their respective value orderings, it also applies when attempting to pool the value orderings of k dimensions. Of course, the second problem may be less of a concern when the individuals agree on which dimensions outweigh the others. Even so, relying on such agreements is dubious: if participants disagree strongly regarding options, they might not agree on relative importance of each dimension of the issue. If the conditions of

(A) The profile of individual preference orderings on a left-to-right continuum based on the increasing expenses of each power plant:



(B) The profile of individual preference orderings on a left-to-right continuum based on the increasing number of jobs created by each power plant:



(c) The collective judgments from employing the Condorcet rule on each of the above single-peaked profiles:

Expenses	Job Creation
nuclear	coal
hydro	hydro
coal	nuclear

FIGURE 3.11: When Dewey, Cheetham, and Howe break the issue of choosing a power plant into two different dimensions—expenses and job creation—each of which has a single-peaked profile, shown in (A) and (B), respectively. For each of these dimensions, the Condorcet rule returns the collective value orderings in (C).

Arrow's theorem apply in cases of collective disagreements between n persons, and this may still be a big if, I see little reason why these conditions do not also apply when there is disagreement over the priority of the k dimensions. Consequently, without an additional account of why the latter is inherently more tractable than the former, I do not find multi-dimensional structuration a plausible way around the impossibility results of social choice theory.

3.6 Responsive Impartiality for Pooling Functions

The conditions of weak Pareto ($\mathcal{P}$) and non-Dictatorship ($\mathcal{D}$) are probably the least controversial of the conditions for pooling functions. Certainly they have generated the least amount of discussion in commentaries on the impossibility theorems. Even so, it is worth keeping in mind that they are thought so unassailable, especially when applied to the decisions of the liberal state, because they jointly make up the basic notion of responsive impartiality. $\mathcal{P}$ ensures responsiveness to consensus while $\mathcal{D}$ protects against the grossest form of partiality. A violation of $\mathcal{P}$ seems simply perverse: why follow a procedure that may judge option y as more valuable than option x even when everyone actually agrees that x is superior than y ? Cases of unanimous agreement should only be overturned when someone changes her mind; that is, when the unanimity between the individual orderings is broken.

One possible rejection of $\mathcal{D}$ might appeal to the idea that often decisions in the liberal state are made by one 'decider'—the state's president, for instance—and not by committee. Even so, in the liberal state, executive decisions are still supposed to be responsively impartial concerning the reasonable views endorsed by the state's constitutive members and not simply partial to the views of whoever happens to control the strings of state power. This raises once again the implication that the results of social choice apply even when certain individuals are making a decision. In the case of the decider in the liberal state, this person is still supposed to remain both rational and responsively impartial in turning several value orderings into a single, all-things-considered judgment. The pooling process may just take place in that person's mind, instead of in a more formal public arena, but that is hardly sufficient to suggest that the internal process should be held to different standards than the external ones. So while the existence of a decider for the liberal state may often be, strictly speaking, true, it is certainly objectionable when that decider is completely insensitive to the value orderings of those who are effected by her decisions. At that point she violates $\mathcal{D}$ and becomes a true enemy of the liberal state.⁵²

⁵²In his *Considerations on Representative Government* (1861), John Stuart Mill expresses a similar

While $(\mathcal{P})$ and $(\mathcal{D})$ may otherwise be non-controversial, the opposite holds for independence of irrelevant alternatives ($\mathcal{IIA}$). Given its formal complexity and sweeping implications, it should not be too surprising that $\mathcal{IIA}$ is usually the locus of criticism against the implications of Kenneth Arrow's theorem. Many of these professed criticisms of $\mathcal{IIA}$, however, are actually directed at a different condition:

Relational Independence ($\mathcal{RJ}$): Consider any pair of options x and y , along with profiles $\mathcal{R}$ and $\mathcal{R}'$. Let R and R' be the respective evaluations returned by the pooling function (i.e., $F(\mathcal{R}) = R$ and $F(\mathcal{R}') = R'$). If both profiles agree with respect to x and y (i.e., for each i , $xR_i y$ holds if and only if $xR'_i y$ holds), then both R and R' must agree with respect to x and y as well.⁵³

According to $\mathcal{RJ}$, a pooling function can only judge a pair of options relative to each other, and only do so based on how each individual value ordering judges that particular pair. Comparisons concerning any other options, regardless of whether these options are on the agenda A or not, are irrelevant. While this is a much stronger condition than $\mathcal{IIA}$, it turns out that given social rationality ($\mathcal{SR}$) and the definition of optimization for $C(A, R)$, $\mathcal{IIA}$ is equivalent to $\mathcal{RJ}$.

Since $\mathcal{RJ}$ is often confused with $\mathcal{IIA}$, it is important to note their distinction.⁵⁴ One way to distinguish them is to notice that $\mathcal{IIA}$ ensures a measure of *inter-profile* consistency: once the options on the agenda A are fixed, the collectively optimal option(s) may change only when individual judgments concerning the options on agenda also change. So options not in A may change their position relative to those

sentiment:

All government must displease many persons. . . . What is the monarch to do when these unfavorable opinions happen to be in the majority? Is he to alter his course? Is he to defer to the nation? If so, he is no longer a despot. [Chapter III in volume 19 of John Stuart Mill, *Collected Works of John Stuart Mill*, ed. by J. M. Robson (Toronto: University of Toronto Press, 1963–1991), p. 402]

⁵³This is condition I^* in Amartya K. Sen, 'Social Choice Theory: A Re-Examination', *Econometrica* 45, 1 (January 1977): 53–89, p. 59.

⁵⁴For instance, Michael Dummett, *Voting Procedures* (Oxford: Clarendon Press, 1984), p. 52; and Dryzek and List, 'Social Choice Theory and Deliberative Democracy: A Reconciliation', p. 4; mistakenly present $\mathcal{RJ}$ as 'Independence of Irrelevant Alternatives'. In addition, an entire chapter in Gerry Mackie, *Democracy Defended* (Cambridge: Cambridge University Press, 2003), pp. 123–157, to address 'Independence of Irrelevant Alternatives' and then proceeds to treat $\mathcal{IIA}$ and $\mathcal{RJ}$ interchangeably without apparently recognizing the difference between the two. Even the standard textbook in public choice theory, Mueller, *Public Choice III*, p. 583, relegates the actual statement of $\mathcal{IIA}$ to a footnote while presenting $\mathcal{RJ}$ as 'Independence of Irrelevant Alternatives' in the main text. For attempts to clarify these confusions see Paramesh Ray, 'Independence of Irrelevant Alternatives', *Econometrica* 41, 5 (September 1973): 987–991; Sen, 'Social Choice Theory: A Re-Examination', p. 59; and Georges Bordes and Nicolaus Tideman, 'Independence of Irrelevant Alternatives in the Theory of Voting', *Theory and Decision* 39, 2 (March 1991): 163–186.

options in A , but such changes do not influence the outcome. Simply put, IJA requires that once the agenda options are fixed, a voting rule should only require the individual evaluations concerning these options. No further information is required.

Both the Condorcet and narrow Borda rules satisfy this demand. The Condorcet rule does so because the particular contents of A are irrelevant for determining the value ordering R . The narrow Borda rule also satisfies inter-profile consistency because it determines the ordering R only after the contents of A have been set. Changing the relative rankings of options off the agenda has no effect on the outcome of the narrow Borda rule. An example of a method that does violating IJA in this way is an alternative formulation of the Borda rule:

The Broad Borda Rule: Say there are an m number of potential options in O . For each top position (relative to all the options in O) held in an R_i , an option receives $m - 1$ points; for each second place position held (relative to those in O), it gets $m - 2$ points; . . . ; for each bottom position held (only relative to those in O), it gets 0 points. xRy if and only if x does not score lower than y .⁵⁵

Unlike the narrow Borda rule, the broad version scores alternatives based on the pairwise comparisons concerning *all* the potential options, including those not on the agenda A . Unlike its narrower cousin, this rule satisfies SR : the scores for all the options are computed the same, irrespective of any changes in the agenda options. Hence, the same weak ordering of options is always returned. This does come at the price of satisfying IJA , however.

To see the problem with the broad Borda rule, consider the two profiles in figure 3.12. Both of these profiles agree with respect to how Argue and Phibbs each order coal and nuclear power—in both profiles Argue orders coal power over nuclear power, and in both Phibbs orders them the other way. But when coal and nuclear power are the only agenda options, the broad Borda rule determines that nuclear power is optimal when given the first profile and yet has coal power as optimal for the second profile.⁵⁶ This is a violation of inter-profile consistency and IJA .

⁵⁵The distinction between the 'Narrow' and 'Broad' Borda Counts was not made by Borda himself, so it is unclear which version he would advocate. For more on this distinction, see Sen, 'Social Choice Theory: A Re-Examination', pp. 78–80. Bordes and Tideman, 'Independence of Irrelevant Alternatives in the Theory of Voting', pp. 178–181, also discuss these rules, which are there called the 'Local' and 'Global' rank-order methods respectively.

⁵⁶In the first profile, nuclear power scores $1 + 2 = 3$ points over coal power's score of $2 + 0 = 2$ points, making it optimal. In the second, coal power is now optimal, scoring $2 + 0 = 2$ points over nuclear power's $0 + 1 = 1$ point.

	$\mathcal{R}_*$		$\mathcal{R}^*$	
	Argue	Phibbs	Argue	Phibbs
1 st	coal	nuclear	coal	hydro
2 nd	nuclear	hydro	hydro	nuclear
3 rd	hydro	coal	nuclear	coal

Broad Borda on Profile . . .		
	$\mathcal{R}_*$	$\mathcal{R}^*$
1 st	nuclear	hydro
2 nd	coal	coal
3 rd	hydro	nuclear

FIGURE 3.12: Two different profiles of individual value orderings representing alternative ways in which Argue and Phibbs could rank coal power, nuclear power, and hydropower. Using the broad Borda rule on each profile leads to the value orderings shown above.

Returning now to $\mathcal{RJ}$, it also demands inter-profile consistency. With a fixed agenda, how two options are collectively ranked relative to each other may only change when individual evaluations concerning those two options change. The example of Argue and Phibbs is therefore a clear violation of $\mathcal{RJ}$. While Argue and Phibbs's respective orderings do not change concerning the relative ranking of coal power versus hydropower across the two profiles, these profiles lead to collective evaluations that do rank coal power and hydropower differently. Even so, $\mathcal{RJ}$ goes further than $\mathcal{JJA}$ by also requiring *inter-agenda* consistency: even as other options are added to or removed from the agenda, how two options are collectively ranked relative to each other may only change when individual judgments concerning these two options change. Given its definition, $\mathcal{JJA}$ alone does not require this because it only involves a fixed agenda and not one that varies. As a result, $\mathcal{JJA}$ does not demand inter-agenda consistency.⁵⁷ This important difference between $\mathcal{JJA}$ and $\mathcal{RJ}$ is seen with the narrow Borda Rule. As already seen, it satisfies $\mathcal{JJA}$, but the example from figure 3.8 reveals that it violates $\mathcal{RJ}$. In that example, the collective ranking of coal power in comparison to nuclear power depends on whether solar power is on the agenda; it does not depend solely on how the group's members rank the two options relative to each other.

Insofar as it ensures inter-agenda consistency, $\mathcal{JJA}$ I believe that it makes a realistic informational demands for the liberal state. The state's constitutive

⁵⁷I take the distinction between inter-profile and inter-agenda consistency from Aldred, 'Social Choice Theory and Deliberative Democracy: A Comment', pp. 750–751. They are respectively known as 'interprofile' and 'intraprofile' consistency in Bordes and Tideman, 'Independence of Irrelevant Alternatives in the Theory of Voting', p. 170.

members need only present their judgments concerning the options on the agenda. In virtually any decision, there exist a vast number of potential options while, for whatever reason, very few of them actually end up on the agenda. So, for example, when determining which power plant to construct, the community need only consider those power plants that are currently available for construction, affordable given budgetary constraints, and whatever other factors may be relevant when selecting a power plant. It would be bizarre to solicit judgments concerning every conceivable power source—a task that would presumably involve a vast number of comparisons.

This concern raises another problem with the broad Borda rule and provides another reason for inter-agenda consistency. A method that flouts inter-agenda consistency like the broad Borda rule faces instability because the precise composition of the set of potential options is necessary for decision making, despite the fact that some potential options are not included on the agenda. This problem is exacerbated by the fact that, as in the case of constructing a power plant, it may often be extremely difficult to ascertain the precise content of the set of potential options but easy to determine those power plants that the community is actually committed to develop. If *JJA* is satisfied, ignorance of all the potential options off the agenda is not a problem. The same optimal options are consistently identified when presented with the same agenda. Of course, more options may later be added or removed from the agenda, and this may certainly alter the optimal options. But this has nothing to do with *JJA*, as it only applies when the agenda is finally fixed. Rather, as we saw with the narrow Borda rule, the problem may be trying to satisfy *SR* when agendas change.

In addition to ensuring inter-profile consistency, *JJA* also mandates an important measure of responsiveness as well by requiring that a pooling function always returns the same optimal option(s) when presented with the same profile of individual value orderings. The idea is expressed by the adage that consistency involves ‘like cases be treated alike’, where like cases according to *JJA* are those where the individuals involved have the exact same ordering of options that are on the agenda *A*. Different cases are those where these rankings differ in some way. In other words, *JJA* ensures that a decision is only sensitive to differences in the individual value orderings and nothing else. This demand immediately eliminates two rules with stochastic properties:

Queen for a Day: The collective value ordering *R* is selected at random from the profile of individual value orderings *R*.

The Imposition of the Day: The collective value ordering *R* is selected at random from all the logically possible value orderings of the options, irrespective of whether anyone actually has that value ordering or not.

Both schemes violate $\mathcal{JJA}$ because when two profiles are exactly the same, each method may return two different collective value orderings with different optimal options. In addition, queen for a day also seems precisely the sort of approach that is against the spirit, if not necessarily the rule, of $\mathcal{D}$.⁵⁸ Although one could imagine a sequence of social choices governed by take-turns dictatorships, this hardly seems an attractive solution either. Meanwhile, the imposition of the day is not attractive as a rule either because it violates $\mathcal{P}$; it is possible for it to select a value ordering that overturns a unanimous agreement concerning some pair of options.

Regardless, similar violations of $\mathcal{JJA}$ hold for virtually any other method that employs stochastic devices. The concern with such methods is that they explicitly make arbitrary decisions insofar as they include elements of chance, and so lack a certain amount of insensitivity to the judgments of the individuals involved. This might not be thought so bad, however. Randomly picking an option seems like a good idea, for instance, if a society is equally and intractably divided between two options. Unfortunately, this does not escape the implications of the impossibility results. Using chance as a tie-breaking device for ultimately picking between multiple top-ranked options does not refute the implications of Arrow's theorem because $\mathcal{SR}$ allows for indifference between options. That is, the results of the theorem still apply to the process used to generate the ordering containing those two top-ranked options. So a violation of one of the conditions is necessary, in order to generate the ordering, even before the seemingly innocuous stochastic device is employed.

Together inter-agenda consistency and this measure of responsive impartiality lead $\mathcal{JJA}$ to require that a collective decision to only employ the information concerning those options 'on the table'. For these reasons, Arrow and others maintain that most 'real-world' voting schemes presume $\mathcal{JJA}$.⁵⁹ Insofar as this is true, I believe that $\mathcal{JJA}$ gets things right. Even so, Arrow's claim might overstate the case for the $\mathcal{JJA}$ because this condition has far more controversial consequences that cannot be ignored. This is because $\mathcal{JJA}$ rules out methods that seek to employ more information than contained in a profile of weak orderings. In particular, $\mathcal{JJA}$ rules out interpersonal value comparisons that might be relevant for decision making by the liberal state.

⁵⁸Queen for a day satisfies $\mathcal{D}$ because it does not necessarily always pick the exact same person as the 'queen'. It is a sort of rotating dictatorship. $\mathcal{D}$ is only against a non-rotating, permanent dictatorship.

⁵⁹Arrow, *Social Choice and Individual Values*, p. 110. Also see Bordes and Tideman, 'Independence of Irrelevant Alternatives in the Theory of Voting', pp. 168–169, 181.

3.6.1 Cardinal Valuations and Interpersonal Comparisons

There are two potential sources of information that Kenneth Arrow's formulation of independence of irrelevant alternatives (*IIA*) has a collective decision effectively ignore: cardinal value assessments and interpersonal comparisons. Recall that I earlier noted (in section 3.1), that individual judgments may contain information concerning the magnitude or intensity of the superiority of one option over another. In that case, an individual may be said to have a value function v_i underlying their value ordering R_i . So in figure 3.2, v_1 and v_2 express the same ordering of the options, but express different magnitudes separating them. According to v_1 , for instance, the magnitude separating options x and z ten-ninths (a little over one) greater than that separating y and z , whereas in v_2 that first magnitude is a full ten times bigger than the second. A focus on value orderings ignores such information.

While I am not suggesting that every value ordering should be derivable from a value function, such cardinal information, when it is available, might be useful for making collective decision. Hence, a pooling function can be augmented into a cardinal pooling function, which is any process taking a profile of individual value functions $\mathcal{V} = \langle v_1, v_2, \dots, v_n \rangle$ and returning a single collective ordering R with which to use for determining the optimal option(s).⁶⁰ Formally speaking, a cardinal pooling function is any function F_C of the form $F_C(\mathcal{V}) = R$.

Supplanting value orderings with value functions in this way does not violate individual rationality (*IR*) because a value function trivially generates a weak ordering of the options.⁶¹ However, the efficacy of cardinal information is diminished by *IIA*. For instance, consider the two alternative profiles of individual value functions for Tipper and Albert in figure 3.13. Since both Tipper and Albert's respective value orderings do not change across these two profiles, *IIA* requires that both profiles should lead to the same collective decision. This might seem troubling, for nuclear power appears to have worse assessments in $\mathcal{V}_A$ than it does in $\mathcal{V}_B$. Some methods might think this is a relevant difference worth considering, but *IIA* holds that it is irrelevant and must be ignored.

The second type of information that *IIA* denies concerns interpersonal comparisons. These types of comparisons differentiate between how one person fares under one option to how another person fares under another option. So, for instance, if the profile of individual value functions $\mathcal{V}_A$ in figure 3.13 reflects inter-

⁶⁰Since all optimization requires is a ranking of the options, in order to select the top ranked one(s), it is sufficient for a cardinal pooling function to return a single weak ordering. Returning a value function is not strictly necessary.

⁶¹Recall from section 3.1 that for all options x and y , xR_iy holds if $v_i(x) \geq v_i(y)$. Mackie, *Democracy Defended*, p. 123, apparently misses this point, where he insists that it is Arrow's definition of a pooling function (social welfare function) that prohibits utility cardinal inputs.

	$\mathcal{V}_A$		$\mathcal{V}_B$	
	Tipper	Albert	Tipper	Albert
coal	1.0	0.0	1.0	0.6
nuclear	0.3	0.3	0.5	0.8
hydro	0.0	1.0	0.0	1.0

FIGURE 3.13: *Two profiles of value functions representing alternative ways in which Albert and Tipper could cardinally assess a decision concerning coal power, nuclear power, and hydropower.*

personal information, it is meaningful to say, for instance, that Tipper's situation under coal power is superior to Albert's under nuclear power, or that both Tipper and Albert have equally valuable situations under nuclear power. If cardinal information is meaningful here, it is then possible to make interpersonal comparisons that also specify the magnitudes of these assessments. In comparing profiles $\mathcal{V}_A$ and $\mathcal{V}_B$ with respect to interpersonal assessments, they seem very different. According to $\mathcal{V}_A$, Tipper's situation under nuclear power is superior to Albert's under coal power, but the opposite judgment is made by $\mathcal{V}_B$. Regardless, $\mathcal{IA}$ holds that both profiles must lead to the same collective decision and so these differences must be ignored.

Even so, there are two possible rules for cardinal pooling functions that do assume that the information from cardinal valuations and interpersonal comparisons are relevant for making a collective decision:

The Summation Rule: Each potential option in O receives a score calculated by taking the sum of the values it receives according to each individual value function in $\mathcal{V}$:

$$s_S(x) = \sum_{i=1}^n v_i(x), \text{ for each option } x.$$

xRy if and only if x does not score lower than y (i.e., $s_S(x) \geq s_S(y)$).

The Maximin Rule: Each potential option in O receives a score calculated by taking the minimum value it receives across all the individual value function in $\mathcal{V}$:

$$s_M(x) = \min_{i=1}^n [v_i(x)], \text{ for each option } x.$$

xRy if and only if x does not score lower than y (i.e., $s_M(x) \geq s_M(y)$).

Summation assesses the options based on the sum that each option receives from the value functions. So when given profile $\mathcal{V}_A$ from figure 3.13, this rule judges

that both coal power and hydropower are the optimal options.⁶² Given its different contents, profile $\mathcal{V}_B$ is treated differently by summation, for in that case only coal power is optimal.⁶³ Maximin scores an option based on the worst score it receives across all the value functions, leading it to evaluate options rather differently than summation. For instance, the optimal option according to maximin for $\mathcal{V}_A$ is nuclear power, while the optimal option for $\mathcal{V}_B$ is coal power.⁶⁴ In both cases, each rule make distinctions that $\mathcal{JJA}$ prohibits.

Once more, I am not suggesting that either cardinal or interpersonal information is always available—perhaps the numbers across the two profiles in figure 3.13 are ultimately meaningless and ought to be ignored. The problem with $\mathcal{JJA}$ is that it rejects these types of information a priori. This is surely too hasty; when such information *is* available, it certainly appears quite reasonable to use it. To this end, it is useful to break $\mathcal{JJA}$ down into the two separate conditions it enforces.

The first of these conditions preserves the measures of inter-agenda consistency and responsive impartiality discussed earlier:

Independence from Irrelevant Alternatives ($\mathcal{JJA}_C$): Consider profiles $\mathcal{V}$ and $\mathcal{V}'$.

Let R and R' be the respective collective value orderings returned by the cardinal pooling function (i.e., $F_C(\mathcal{V}) = R$ and $F_C(\mathcal{V}') = R'$). When given an agenda of available options A , if both profiles agree with respect to this agenda (i.e., for each i , and all x in A , then $v_i(x) = v'_i(x)$), then $C(A, R) = C(A, R')$.⁶⁵

This condition says that when two profiles of value functions have identical numbers for each corresponding pair of value functions when given an agenda, then the collective decision must be the same. But if these numbers differ anywhere, then the two profiles may result in a different collective decision. As a result, profiles $\mathcal{V}_A$ and $\mathcal{V}_B$ from figure 3.13 may be treated differently according to $\mathcal{JJA}_C$. Furthermore, both the summation and maximin satisfy this condition.

The second condition denies employing both cardinal information and interpersonal comparisons:

Ordinal Non-Comparability ($\mathcal{ONC}_C$): Consider profiles $\mathcal{V}$ and $\mathcal{V}'$. Let R and R' be the respective collective value orderings returned by the cardinal pooling function (i.e., $F_C(\mathcal{V}) = R$ and $F_C(\mathcal{V}') = R'$). When given an agenda of available options A , if both profiles agree across their corresponding

⁶² $s_S(\text{coal}) = 1.0 + 0.0 = 1.0$, $s_S(\text{nuclear}) = 0.3 + 0.3 = 0.6$, and $s_S(\text{hydro}) = 0.0 + 1.0 = 1.0$.

⁶³ $s_S(\text{coal}) = 1.0 + 0.6 = 1.6$, $s_S(\text{nuclear}) = 0.5 + 0.8 = 1.3$, and $s_S(\text{hydro}) = 0.0 + 1.0 = 1.0$.

⁶⁴For $\mathcal{V}_A$, $s_M(\text{coal}) = 0.0$, $s_M(\text{nuclear}) = 0.3$, and $s_M(\text{hydro}) = 0.0$. Meanwhile, for $\mathcal{V}_B$, $s_M(\text{coal}) = 0.6$, $s_M(\text{nuclear}) = 0.5$, and $s_M(\text{hydro}) = 0.0$.

⁶⁵Sen, *Collective Choice and Social Welfare*, p. 129.

orderings with respect to this agenda (i.e., for each i , and all x and y in A , $v_i(x) \geq v_i(y)$ holds if and only if $v'_i(x) \geq v'_i(y)$ holds), then $C(A, R) = C(A, R')$.⁶⁶

In other words, ONC_C requires that when the corresponding value functions across two profiles both order the options the same way, then the collective decision must be the same for both profiles. It is then a violation of this condition to treat the profiles from figure 3.13 differently, despite the fact that these matrices have different numbers. Any potential cardinal and interpersonal information revealed in these matrices must be ignored, and so both summation and maximin violate this condition.

Converting the remainder of Arrow's conditions into the framework for cardinal pooling functions is straightforward:

Individual Rationality ($\mathcal{IR}_C$): The cardinal pooling function only accepts profiles $\mathcal{V}$ where each individual value function v_i in $\mathcal{V}$ generates a weak ordering of the potential options.

Social Rationality ($\mathcal{SR}_C$): The cardinal pooling function always returns an evaluation R that is itself a weak ordering of the potential options.

Weak Pareto Principle ($\mathcal{P}_C$): If for all v_i in $\mathcal{V}$, $v_i(x) > v_i(y)$ holds for potential options x and y , then the cardinal pooling function must return a collective value ordering R satisfying xPy .

Non-Dictatorship ($\mathcal{D}_C$): The cardinal pooling function cannot be dictatorial. That is, there is no individual d in N , such that for all profiles $\mathcal{V}$ and all potential options x and y , whenever $v_d(x) > v_d(y)$ holds the cardinal pooling function always returns a collective value ordering R satisfying xPy .

Unrestricted Domain ($\mathcal{U}_C$): The domain of the cardinal pooling function is the set of all logically possible profiles $\mathcal{V}$.

Since any value function trivially generates a weak ordering of the options, $\mathcal{IR}_C$ is always true, but it serves as a reminder of the assumption concerning the rationality of the individual value orderings. The other conditions are practically identical to their counterparts in Arrow's original framework.

In this new framework, Amartya Sen proves the equivalent of Arrow's theorem:

⁶⁶Kevin W. S. Roberts, 'Interpersonal Comparability and Social Choice Theory', *Review of Economic Studies* 47, 2 (January 1980): 421–439, p. 423. This condition is also known as ON in Claude D'Aspremont and Louis Gevers, 'Equity and the Informational Basis of Collective Choice', *Review of Economic Studies* 44, 2 (June 1977): 199–209, p. 200, and as ordinal invariance in Kenneth J. Arrow, 'Extended Sympathy and the Possibility of Social Choice', *American Economic Review* 67, 1 (February 1977): 219–225, p. 220.

Sen's ONC Impossibility Theorem: There is no non-trivial cardinal pooling function satisfying $\mathcal{IR}_C$, $\mathcal{SR}_C$, $\mathcal{P}_C$, $\mathcal{D}_C$, $\mathcal{U}_C$, $\mathcal{IIA}_C$, and $\mathcal{ONC}_C$.⁶⁷

This reveals more clearly the information constraints that Arrow's original conditions presume concerning ordinal valuations and interpersonal comparisons. When this information is ignored, then there is no method for making collective decisions that are both rational and responsively impartial.

Naturally, a promising approach might be to still block interpersonal information, but allow a scheme to distinguish information concerning the cardinal magnitudes separating each pair of options in each individual value function. This is formalized by the following condition:

Cardinal Non-Comparability ($\mathcal{CNC}_C$): Consider profiles $\mathcal{V}$ and $\mathcal{V}'$. Let R and R' be the respective collective value orderings returned by the cardinal pooling function (i.e., $F_C(\mathcal{V}) = R$ and $F_C(\mathcal{V}') = R'$). When given an agenda of available options A , if for each i , v'_i is a strictly positive affine transformation of v_i relative to those options (i.e., for each i , and all x in A , $v'_i(x) = \alpha_i + (\beta_i \times v_i(x))$, where each α_i and each β_i is a real number with each $\beta_i > 0$), then $C(A, R) = C(A, R')$.⁶⁸

A strictly positive affine transformation of a value function preserves the relative magnitudes between pairs of options.⁶⁹ Therefore, a process satisfying $\mathcal{CNC}_C$ may make different collective decisions for profiles $\mathcal{V}_A$ and $\mathcal{V}_B$ from figure 3.13.

⁶⁷This is an immediate consequence of theorem 8*2 in Sen, *Collective Choice and Social Welfare*, p. 129.

⁶⁸Roberts, 'Interpersonal Comparability and Social Choice Theory', p. 423. This condition is also known as CN in D'Aspremont and Gevers, 'Equity and the Informational Basis of Collective Choice', p. 200.

⁶⁹A strictly positive affine transformation preserves all the ratios between all the pairs of magnitudes separating options. More precisely, if $v'_i(x)$ is a strictly positive affine transformation of $v_i(x)$, then given any options w, x, y , and z , the following ratio remains constant:

$$\frac{v_i(w) - v_i(x)}{v_i(y) - v_i(z)}.$$

So, for instance, Tipper's value function in $\mathcal{V}_A$ from figure 3.13 is not a strictly positive affine transformation of her function in $\mathcal{V}_B$ from that figure because according to the first one, the magnitude between coal power and hydropower is three-and-a-third times bigger than the magnitude between nuclear power and hydropower, whereas it is only two times bigger according to the second one. However, Tipper's value function in $\mathcal{V}_C$ from figure 3.14 is a strictly positive affine transformation of her function in $\mathcal{V}_A$, with $\alpha_i = 1.0$ and $\beta_i = 2.0$ (where i represents the index of Tipper's value function in these profiles). Notice that this means the relative magnitudes between options are the same for both of her value functions. So the magnitude between her evaluations of coal power and hydropower remains three-and-a-third times greater than the magnitude between nuclear power and hydropower according to both of these value functions.

	$\mathcal{V}_C$	
	Tipper	Albert
coal	3.0	0.0
nuclear	1.6	0.3
hydro	1.0	1.0

FIGURE 3.14: A profile of value functions representing an additional way in which Albert and Tipper could cardinaly assess a decision concerning coal power, nuclear power, and hydropower.

Even so, the summation and maximin rules do not satisfy CNC_C . To illustrate, this condition demands that profile $\mathcal{V}_A$ from figure 3.13 and $\mathcal{V}_C$ from figure 3.14 be treated identically.⁷⁰ Summation, however, judges that both coal power and hydropower are optimal for $\mathcal{V}_A$ while also judging that only coal power is optimal for $\mathcal{V}_C$, and so summation violates CNC_C .⁷¹ Similarly, maximin identifies a different optimal option for each profile. According to maximin, nuclear power is optimal for $\mathcal{V}_A$ whereas hydropower is optimal for $\mathcal{V}_C$.⁷² Making such a distinction violates CNC_C .

On reflection it should not really be all that surprising that both summation and maximin cannot satisfy CNC_C , for both rules rely on interpersonal comparisons to make collective judgments. Summation needs interpersonal comparisons so that the units of measurement across the different value functions remain constant. One unit of improvement to one value functions should be the same amount of improvement according to the other ones. Meanwhile, maximin needs interpersonal comparisons in order to identify which assessment gives the lowest value. If the evaluations of one function cannot be compared to those of another, it is meaningless to say one function's assessment of an option is lower than another's. So going from $\mathcal{V}_A$ in figure 3.13 and $\mathcal{V}_C$ in figure 3.14 lays out extremely different interpersonal comparisons. In the first, Tipper under hydropower is inferior to Albert under nuclear power, but in the second reverse this judgment.

Regardless, it might be thought that allowing cardinal information regarding magnitudes, even while denying interpersonal comparisons, might still avoid the implications of Arrow's theorem. It happens that this is not true, as Sen proves:

Sen's CNC Impossibility Theorem: There is no non-trivial cardinal pooling

⁷⁰Albert's value function returns the same values in both profiles. While Tipper's respective functions do return different values, the one in $\mathcal{V}_C$ is a strictly positive affine transformation of the one in $\mathcal{V}_A$ (see n. 69 in this chapter).

⁷¹For $\mathcal{V}_C$, $s_S(\text{coal}) = 0.0$, $s_S(\text{nuclear}) = 0.3$, and $s_S(\text{hydro}) = 1.0$.

⁷²For $\mathcal{V}_C$, $s_M(\text{coal}) = 0.0$, $s_M(\text{nuclear}) = 0.3$, and $s_M(\text{hydro}) = 1.0$.

function satisfying $\mathcal{IR}_C, \mathcal{SR}_C, \mathcal{P}_C, \mathcal{D}_C, \mathcal{U}_C, \mathcal{IIA}_C$, and $\mathcal{CNC}_C$.⁷³

Apart from this result, it even turns out that considering even more precise cardinal information, for instance, allowing a distinction to be made between $\mathcal{V}_A$ from figure 3.13 and $\mathcal{V}_C$ from figure 3.14 is still subject to impossibility. However, introducing interpersonal comparisons is enough to open up possibilities such as the summation and maximin rules.⁷⁴

The problem that emerges with interpersonal comparisons is a serious embarrassment of riches along with what appear to be widespread reasonable disagreements concerning which of these riches the liberal state should seize upon in order to make its decisions. In particular, when it comes to interpersonal comparisons, there are at least three areas of controversy:

- (1) The proper metric to use for interpersonal comparisons,
- (2) The data reported by a given metric, and
- (3) The proper way to use the data for making a collective decision.

The first area of controversy is exemplified by the debate over the space of equality. I believe that the most promising candidates here are John Rawls' theory of 'primary goods' and Sen and Martha Nussbaum's respective accounts of 'functionings and capabilities'.⁷⁵ Unlike subjective notions such as 'happiness', 'pleasure', or 'satisfaction', these standards are non-mentalistic and externally measurable. They are ideal for collective decision making by the liberal state because they provide public measures justifiable to everyone. Even so, I must concede there is a reasonable room for disagreement here, not only between Rawls, Sen, and Nussbaum, but also between them and those advocating measures of equality based upon

⁷³This is another immediate consequence of theorem 8*2 in Sen, *Collective Choice and Social Welfare*, p. 129.

⁷⁴Possibility results concerning these two rules may be found in D'Aspremont and Gevers, 'Equity and the Informational Basis of Collective Choice', as well as Robert Deschamps and Louis Gevers, 'Leximin and Utilitarian Rules: A Joint Characterization', *Journal of Economic Theory* 17, 2 (April 1978): 143–163. A useful resource for comparing different possible rules for cardinal pooling functions may be found in table 1 in Roberts, 'Interpersonal Comparability and Social Choice Theory', p. 437.

⁷⁵John Rawls, 'Social Utility and the Primary Goods', in *Utilitarianism and Beyond*, ed. by Amartya K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982): 159–186, and 'The Priority of Right and Ideas of the Good', *Philosophy and Public Affairs* 17, 4 (Autumn 1988): 251–276; Amartya K. Sen, 'Equality of What?', in *Tanner Lectures on Human Values*, ed. by Sterling M. McMurrin, vol. 1 (Salt Lake City: University of Utah Press, 1980): 195–220, *Commodities and Capabilities*, (Amsterdam: North-Holland, 1985), and *Inequality Reexamined*, (Oxford: Oxford University Press, 1992); and Martha C. Nussbaum, *Sex and Social Justice* (Oxford: Oxford University Press, 1999), 'Women and Equality: The Capabilities Approach', *International Labour Review* 138, 3 (September 1999): 227–245, and *Women and Human Development*, (Cambridge: Cambridge University Press, 2000).

‘welfare’, ‘advantage’, ‘resources’, and even ‘compassion’ and ‘extended sympathy’.⁷⁶ If there is this much debate amongst political theorists, I see little reason why the state’s constitutive members would be more likely to agree on a metric for interpersonal comparisons.

Furthermore, even if such a metric is agreed upon, there may exist reasonable disagreements concerning the data being measured. Jone’s might claim that implementing policy *x* gives her fewer primary goods than Smith receives under it. Smith might counter, that as he sees it, Jone’s is getting more than enough. Recall that Rawls’ very first ‘burden of reason’ (which I presented in section 1.2.2) involves reasonable disagreements over the evidence bearing on a decision.⁷⁷ Consequently, it is acceptable for the state’s constitutive members to disagree over the numerical contents reported by the value functions being used by the liberal state to make a collective decision.

Finally, even if the metric is agreed upon, and there is no controversy concerning the data that the value functions report for that metric, there remains the concern of how exactly to pool those functions into a single, all-things-considered collective judgment. As has already been seen, both summation and maximin are available. In addition, there is reasonable debate concerning other methods as well, concerning methods that are ‘egalitarian’, ‘prioritarian’, and ‘sufficiency based’.⁷⁸ While my own sympathies here lie with sufficiency methods, I must once

⁷⁶For an advocate of welfare, see Richard J. Arneson, ‘Equality and Equal Opportunity for Welfare’, *Philosophical Studies* 56, 1 (May 1989): 77–93. Gerald A. Cohen, ‘On the Currency of Egalitarian Justice’, *Ethics* 99, 4 (July 1989): 906–944, advocates advantage; Ronald Dworkin, ‘What is Equality? – Part 2: Equality of Resources’, *Philosophy and Public Affairs* 10, 4 (Autumn 1981): 283–345, advocates resources; Roger Crisp, ‘Equality, Priority, and Compassion’, *Ethics* 113, 4 (July 2003): 745–763, advocates compassion; and even Arrow, ‘Extended Sympathy and the Possibility of Social Choice’, finally acknowledges an interpersonal metric of sympathy. For reasons best articulated by both Rawls, Sen and Nussbaum, I do not find these spaces appealing for decision making by the liberal state, but I do not believe that their respective defenders are being unreasonable in so advocating them.

⁷⁷John Rawls, *Political Liberalism* (New York: Columbia University Press, 1993), p. 56.

⁷⁸Amartya K. Sen, ‘Interpersonal Aggregation and Partial Comparability’, *Econometrica* 38, 3 (May 1970): 393–409, ‘Interpersonal Comparison and Partial Comparability: A Correction’, *Econometrica* 40, 5 (September 1972): 959, and ‘Interpersonal Comparisons of Welfare’, in *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, ed. by Michael J. Boskin (New York: Academic Press, 1979): 183–201; and Roberts, ‘Interpersonal Comparability and Social Choice Theory’, and ‘Possibility Theorems with Interpersonally Comparable Welfare Levels’, *Review of Economic Studies* 47, 2 (January 1980): 409–420, canvass a wide variety of options. The debate concerning egalitarian versus prioritarian versus sufficiency methods is seen in Jeremy Waldron, ‘John Rawls and the Social Minimum’, *Journal of Applied Philosophy* 3, 1 (March 1986): 21–33; Harry Frankfurt, ‘Equality as a Moral Ideal’, *Ethics* 98, 1 (October 1987): 21–43; Elizabeth S. Anderson, ‘What Is the Point of Equality?’, *Ethics* 109, 2 (January 1999): 287–337; Derek Parfit, ‘Equality or Priority’, in *The Ideal of Equality*, ed. by Matthew Clayton and Andrew Williams (London: Macmillan, 2000):

again concede that I cannot claim that advocates of the other methods are being unreasonable.⁷⁹ So yet again, the constitutive members of the liberal state may reasonably disagree over the choice of pooling method.

Putting these three issues together, relying on interpersonal comparisons does not provide much comfort. Nor does it necessarily avoid the impossibility results. For instance, say primary goods with data set D_1 under maximin yields value ordering R_1 , while functionings with data set D_2 under summation leads to a different value ordering R_2 . Given the various permutations of the three areas of disagreement over interpersonal comparisons, a large profile of reasonable value orderings could be produced. Without agreement on which ordering represents the proper one for the liberal state to employ—and the fact of reasonable pluralism allows for reasonable disagreement here, brings us back once again to the implications of the impossibility results. As I maintained regarding multidimensional structuration (in section 3.5.1), these results apply in all instances of pooling a collection of multiple value orderings. In this case, the value orderings are those generated by different understandings of interpersonal comparisons.

In light of these concerns, interpersonal comparisons certainly appear very promising, but the fact of reasonable pluralism and the burdens of reason bring us back to the dismal results of social choice theory.

3.7 Implications for Liberal Democracy

For the most part, political scientists and economists draw rather stark conclusions from the impossibility results of social choice theory. This position is best exemplified by William Riker, the founder of the influential Rochester School of political science, who maintain that Kenneth Arrow's theorem demonstrates that democratic institutions unavoidably make arbitrary and meaningless decisions. He points out that the disagreement between the Condorcet rule, the various Borda rules, and the other decision making rules over outcomes (recall this was seen with the profile in figure 3.7) illustrates why collective decision making is ultimately arbitrary. The particular method employed significantly determines the outcome, and since any such method must violate one of the theorem's conditions, there is apparently nothing privileging one method over the other.

On this matter, Riker is almost shrill:

So long as a society preserves democratic institutions, its members

81–125; Crisp, 'Equality, Priority, and Compassion'; and Larry S. Temkin, *Inequality* (Oxford: Oxford University Press, 1996), and 'Egalitarianism Defended', *Ethics* 113, 4 (July 2003): 764–782.

⁷⁹I believe Waldron, 'John Rawls and the Social Minimum', in particular, provides a compelling argument for a social minimum.

can expect that some of their social choices will be unordered or inconsistent. And when this is true, no meaningful choice can be made. If y is in fact chosen—given the mechanism of choice and the profile of individual valuations—then to say that [some other] x is best or right or more desired is probably false. But it would be equally false to say that y is best or right or most desired. And in that sense, the choice lacks meaning.⁸⁰

Despite this problem, Riker does believe that the liberal state should be democratic, though only in a minimal sense.⁸¹ Minimal democracy emphasizes that the liberal state take the form of a strict constitutional republic with features such as multicameralism, federalism, and a balance of powers, where elected representatives are society's decision makers. Popular participation is restricted to selecting these representatives. By Riker's account, elections allow society's members to protect themselves from tyrants by controlling elected officials and their policies with the threat of removal from office. Popular participation therefore ensures the possibility of removing oppressive rulers from power—regular elections provide this opportunity and nothing more. Considering Arrow's theorem, Riker rejects other forms of participation, such as referenda and public opinion polls, because they only produce 'inconsistent and bizarre legislation' that 'puts democracy at risk'.⁸²

Critics of this defense of minimal democracy typically respond to Riker's position with something like the following:

⁸⁰Riker, *Liberalism Against Populism*, p. 136. In light of other social choice results, especially a theorem proved independently by Allan Gibbard, 'Manipulation of Voting Schemes: A General Result', *Econometrica* 41, 4 (July 1973): 587–601, and Mark A. Satterthwaite, 'Strategy-Proofness and Arrows Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions', *Journal of Economic Theory* 10, 2 (1975): 187–217, that purports to show that no pooling function is immune to strategic manipulation, Riker goes even further:

[A social choice] may consist of the amalgamation of the true tastes of the majority . . . or they may consist simply of the tastes of some people . . . who are skillful or lucky manipulators. If we assume that social choices are often of the latter, they may consist of what the manipulators truly want, or they may be an accidental amalgamation of what the manipulators (perhaps unintentionally) happen to produce. Furthermore, since we can by observation know only expressed values (never true values), we can never be sure, for any particular social choice, which of these possible interpretations is correct. [Riker, *Liberalism Against Populism*, pp. 167–168]

While this concern with eliciting 'honest' preferences from the state's constitutive members is a serious one, I do not pursue it in this dissertation. Instead I continue to focus on the problem of pooling value orderings, even when they are *honest*, raised by Arrow's theorem.

⁸¹ibid., pp. 241–246. What I call minimal, Riker calls 'liberal'.

⁸²ibid., pp. 251, 252.

Nonreasoned removal from office is precisely what follows if Riker is correct in interpreting the instability results of social choice theory as demonstrating the meaninglessness of voting. If outcomes are arbitrarily connected to the preferences of the electorate, we cannot infer from the removal from office that an officeholder's conduct was in fact disapproved of by the voters. This is hardly the ideal of officeholders being put at risk by elections that we associate with [minimal] democracy.⁸³

Given that Riker's theory relies on pooling information when electing representatives, his own claims concerning the impossibility results are then redirected towards his own arguments supporting a minimal form of democracy. However, this criticism misses the point: Riker is well aware of how his arguments apply to minimal democracy.⁸⁴ Minimalism is a sound theory since it only claims that voting creates the *possibility* of removing tyrants; nowhere does Riker suggest that minimal democracy must *always* remove them. Nor does he suggest that an unoffending ruler must always stay in office. All that matters is that regular elections ensure the possibility, not the guarantee, of a public veto on tyranny always remains.

Nevertheless, two deeper problems face minimal democracy. First of all, Riker is clearly ambivalent towards the reliable removal of tyrants through public participation. Furthermore, he claims that such participation 'need not be fair or just' in its outcomes.⁸⁵ This naturally leads to the concern that he regards popular participation and democratic control as merely symbolic gestures lacking all substance. Given Riker's case against pooling value orderings, it seems that all democratic theorists ought to recognize that public participation is either utterly undemocratic or unavoidably irrational. In fact, regular selection of representatives by lot is an acceptable alternative to elections. That is, random allocation of offices certainly ensures the possibility of removing a tyrant. Depending on how often aggregation results in irrational decisions, selection by lots may even be more effective than elections in this regard. Popular participation in government, an essential aspect of democracy, becomes severely limited or, depending on how restrictive the conditions for being selected as a representative, all but eliminated. Riker has eroded away the possibility of rational collective decision making and so, to a certain extent, democracy for minimal democracy.

⁸³Jules Coleman and John Ferejohn, 'Democracy and Social Choice', *Ethics* 97, 1 (October 1986): 6–25, p. 22.

⁸⁴For instance, see Riker, *Liberalism Against Populism*, pp. 242–244. One commentator who does realize this is Joshua Cohen, 'An Epistemic Conception of Democracy', *Ethics* 97, 1 (October 1986): 26–38, pp. 30–31.

⁸⁵Riker, *Liberalism Against Populism*, p. 244.

Even more serious, however, is the fact that the same problems with collective decision making facing popular participation inevitably reoccur within the decision-making processes of the representatives. Certainly fewer individuals are involved, but Arrow's theorem nonetheless applies, implying the inevitability of arbitrary and meaningless decisions by legislators. That is, unless there is only one legislative branch with one representative (the 'dictator'), Arrow's theorem applies. With multicameralism and multiple representatives, there is presumably collective decision making within each branch of the legislature and then further collective decision making between these branches before a final option is chosen. Riker has not removed the problem of collective decision making so much as he has merely moved from the populace as a whole to the members of the legislature or legislatures. In responding to this concern, Riker might appeal to the idea that the representatives do not face the same intractable pluralism and conflict as their constituents. This suggests that unrestricted domain ($\mathcal{U}$) need not apply to their decisions. For the same reasons I had in response to the arguments of John Dryzek and Christian List in section 3.5 concerning domain restriction, I see little reason to believe that this is true. Consequently an alternative approach is needed if liberal democracy is to emerge from the shadows of social choice theory. I believe that while the default rationality thesis is attractive in its parsimony, there exist good reasons for abandoning it and thereby undermine the dire implications of the impossibility results for liberal democracy. I now turn to this issue.

CHAPTER FOUR

Reconciling Asymmetries in Reasoning

KENNETH Arrow's theorem and the theorem due to Teddy Seidenfeld, Joseph Kadane, and Mark Schervish suggest there exist serious limitations for how the liberal state may rationally derive a collective, all-things-considered assessment of options from the rational individual judgments of those same options by the group's constitutive members. Both theorems suppose that a process for transforming individual assessments into a social choice ought to make some decision no matter what reasonable views the individuals may possess, so long as these views meet what appear to be rather minimal standards. In particular, the rationality hypotheses of these theorems maintain that each individual make value judgments conforming to the default rationality thesis. This means that these judgments should comprise a weak ordering of the options satisfying reflexivity, transitivity, and completeness. The problem raised by the social choice theorems is that even when all of the liberal state's constitutive member's possess rational judgments in this sense, it is impossible to avoid all instances where the commitment to responsive impartiality leads to irrational collective judgments. So appealing to a system of democratic institutions for decision making is not enough to ease the tension (from section 1.4) between responsive impartiality and state agency.

As I mentioned in the previous chapter, this issue leads many in political science and economics to conclude that it is impossible for the liberal state to remain responsively impartial in the face of societal pluralism. They tend to believe that the liberal state should either recede into the more minimal role of providing

collective security, or that the state should give up pretensions to impartiality and enforce a comprehensive doctrine of its own. However, Arrow's theorem has also, in part, subtly influenced these disciplines in another, often overlooked, way. It has caused many political scientists and economists to believe, whether explicitly or not, that it is either unnecessary or even logically inconsistent to assess collective decisions by norms of rationality. State decisions cannot be judged by the same norms as those made by individual persons; only the judgments of an individual ought to conform to the default rationality thesis.

However, the idea that the judgments of individual persons are, in fact, so nicely arranged has been rather vividly denied by various thinkers, from Plato to Sigmund Freud to Donald Davidson, and by any number of recent psychological studies.¹ Plato, Freud, and Davidson think that an individual person is more like a community that is already incoherent, at least according to the default rationality thesis, than like a individual with a single, unified, all-things-considered, 'rational' structure of value judgments. That individuals are often irrational according to the default rationality thesis would seem only to make democracy a foolish combination of foolish evaluations. Reason's struggle with the appetite, the Ego clashing with the Id, or conflicting internal processes producing weakness of will, and other supposed irrationalities of judgment, do not seem a likely place to look for a way to neutralize the social choice theorems.

My contention is that, appearances notwithstanding, this *does* point the way for undermining the relevance of these theorems while revealing a more promising way to relieve the deeper tension between responsive impartiality and state agency. I believe that Plato, Freud, and Davidson uncover an important descriptive claim: an individual person often find herself in decision-making situations that resemble those faced by the liberal state. That is, there are fairly routine cases of individual

¹Plato's *Republic* and *Phaedrus* are, of course, explicitly concerned with this. However, Plato believes that while most people have internal conflicts between their rational, appetitive, and spirited parts, a select few—the philosopher-kings—possess an ordered and unified soul. On the other hand, Freud's *Ego and the Id* expresses doubts that such internal harmony is possible for anyone. For more on how Plato's analyses of individuals and groups are two aspects of the same psychological theory, see Jonathan Lear, 'Inside and Outside the *Republic*', in *Plato's Republic: Critical Essays*, ed. by Richard Kraut (Lanham, MD: Rowman and Littlefield, 1997): 61–94. Clark Glymour, 'Freud's Androids', in *The Cambridge Companion to Freud*, ed. by Jerome Neu (Cambridge: Cambridge University Press, 1991): 44–85 argues that Freud anticipates elements of Arrow's theorem. Davidson's representative work on this issue include Donald Davidson, 'How is Weakness of the Will Possible?', in *Moral Concepts*, ed. by Joel Feinberg (Oxford: Oxford University Press, 1969): 93–113, and 'Paradoxes of Irrationality', in *Philosophical Essays on Freud*, ed. by Richard Wollheim and James Hopkins (Cambridge: Cambridge University Press, 1982): 289–305. The *locus classicus* for the general failure of individuals to conform to the default rationality thesis remains Amos Tversky and Daniel Kahneman, 'Judgement Under Uncertainty: Heuristics and Biases', *Science* 185, 4157 (1974): 1124–1131.

decision making that have an analogous structure to those made by the liberal state, whereby a choice must be made in the face of a plurality of incompatible desires or commitments. These thinkers go wrong, however, by suggesting that in such situations a person is unavoidably irrational if she cannot bring order to this plurality and resolve her inner conflicts. Consequently, they ally themselves with the political scientists and economists who maintain that without increased systematicity and conformity to the default rationality thesis, decisions run the risk of being ad hoc and arbitrary. To the contrary, however, I maintain that the normalcy of these situations for the individual person suggests the need for a normative account of decision making under unresolved conflict. If such an account is appealing for guiding the decisions by individuals, I argue that it should be extended to assess those of the liberal state.

So while in Plato's *Republic*, Socrates proposes discovering the organization of the soul through examining the structure of the *polis*, I recommend emulating Socrates in reverse.² Rational deliberation of the liberal state may be better apprehended by examining that of individuals in situations of unresolved conflict. I believe this offers a more promising route than the more common, and ultimately less convincing, ways I addressed in the previous chapter for avoiding the implications of the social choice impossibility results. These less convincing defenses of liberal democracy focus on how group decisions can remain responsively impartial despite the impossibility results, while virtually no liberal democrats examine how individual persons deliberate when making decisions, what norms ought to apply to these deliberations, and determine the extent to which this informs an account for how the liberal state ought to make its decisions.³

In defending this approach, I assume—contrary to many in political science and economics—that the decisions of both individual persons and the liberal state ought to be judged by the same norms of rationality. I call this assumption the symmetry thesis, and I seek in this chapter to defend it and show how it may open up a path for denying the pessimistic implications of the social choice impossibility theorems for the liberal state. I begin by introducing and rejecting what I take to be two common objections against the idea of applying any norms of rationality whatsoever to assess the decisions of the liberal state (section 4.1). This gives me the opportunity to clarify why I believe it is acceptable to treat the liberal state as a rational agent. Following that, I respond to the deeper objection that the individual

²*Republic*, book II [368c–369b] in Plato, *Complete Works*, ed. by John M. Cooper and D. S. Hutchinson (Indianapolis, IN: Hackett, 1997), pp. 1007–1008.

³One notable exception to this, though one going in an extremely different direction than I do, appears in Robert E. Goodin, 'Democratic Deliberation Within', *Philosophy and Public Affairs* 29, 1 (Winter 2000): 79–107, which is expanded in *Reflective Democracy*, (Oxford: Oxford University Press, 2003), pp. 194–225.

and the liberal state possess fundamentally different commitments, which implies that the same norms of rationality should not apply to both (section 4.2). Finally, I conclude with why I believe the symmetry thesis entails abandoning the default rationality thesis, and in doing so, undermines the relevance of the impossible theorems (section 4.3). In the next chapter, I then argue for what norms of rationality I believe should apply to decision making under unresolved conflict, regardless of whether the decision is made by an individual person or the liberal state.

4.1 The Liberal State as Rational Agent

The suspicion against applying the same norms of rationality to the decisions of both individual persons and the state is common in contemporary political science and economics. This suspicion is so persuasive and commonplace that no one in these disciplines seems to believe it even necessary to mount a clear elaboration and defense of it. A useful starting point to understanding the rationale behind it, however, appears in the writings of James Buchanan, who does mount a modest attack on the symmetry thesis.⁴ First, Buchanan holds that it is unnecessary to assess state decisions with the norms of the default rationality thesis. Second, he argues that seeking to assess state decisions in this way presents insurmountable philosophical difficulties. Together these underpin Buchanan's basic denial of the very idea of 'collective rationality'.⁵

Underlying this position is Buchanan's assumption of a strict 'either-or' taxonomy of possible 'philosophical bases': theories are either 'individualistic' in assessing individual persons by the norms of rationality or 'organic' in also doing so for the state.⁶ In individualistic theories, the state 'is represented as the sum of

⁴James M. Buchanan, 'The Pure Theory of Government Finance: A Suggested Approach', *Journal of Political Economy* 57, 6 (December 1949): 496–505, 'Individual Choice in Voting and the Market', *Journal of Political Economy* 62, 4 (August 1954): 334–343, and 'Social Choice, Democracy, and Free Markets', *Journal of Political Economy* 62, 2 (April 1954): 114–123. Buchanan's position is clearly influenced by the methodological individualism of Friedrich A. Hayek, 'Scientism and the Study of Society: Part I', *Economica* 9, 35 (August 1942): 267–291, 'Scientism and the Study of Society: Part II', *Economica* 10, 37 (February 1943): 34–63, and 'Scientism and the Study of Society: Part III', *Economica* 11, 41 (February 1944): 27–39; and Karl Popper, 'The Poverty of Historicism, I', *Economica* 11, 42 (May 1944): 86–103, 'The Poverty of Historicism, II', *Economica* 11, 43 (August 1944): 119–137, and 'The Poverty of Historicism, III', *Economica* 12, 46 (May 1945): 69–89.

⁵According to one commentator, Buchanan holds 'the creed of free trade, limited government, and fiscal conservatism', leading 'Buchanan and his collaborators . . . [to uphold] the premise that any discussion of public interest or social welfare [violates] their commitment to individualistic philosophy', Sonja M. Amadae, *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism* (Chicago: University of Chicago Press, 2003), pp. 135–136.

⁶Buchanan, 'Social Choice, Democracy, and Free Markets', p. 116. Buchanan introduces this

its individual members acting in a collective capacity',⁷ with 'no ends other than those of its individual members and [the state] is not a separate decision-making unit'⁸ and therefore 'no question of social or collective rationality may be raised'.⁹ Buchanan then maintains that, in fact, liberal democratic institutions, as well as those of the free market, have been founded on an individualistic philosophy and not an organic one.¹⁰ Consequently, there is no need to directly assess the rationality or irrationality of the liberal state's decisions. It suffices to instead examine the decisions of the individual persons involved.

Buchanan does not stop there, however. He asserts that organic theories amount to metaphysical hokum:

The mere introduction of the idea of social rationality suggests the fundamental philosophical issues involved. Rationality or irrationality as an attribute of the social group implies the imputation to that group of an organic existence apart from that of its individual components.¹¹

Here I believe Buchanan is claiming that by attempting to assess state decisions by the norms of rationality, organic theories must treat the state as if it were itself a rational agent. Why exactly, this is metaphysically or ontologically problematic, Buchanan does not make more explicit. However, I imagine he is assuming that, unlike the state, an individual person has a unity of body and a unity of consciousness, and that these allow her to provide a rational unity to her own values, commitments, and resultant decisions.¹² It therefore makes sense, according to this line of thought, to treat a person as a rational agent by assessing her decisions. The argument therefore concludes that treating the liberal state as a rational agent rests on nonsensical assumptions about the unity of body and of consciousness, which the state does not possess. I believe this explains why Buchanan believes assessing the rationality of state decisions is involves problematic assumptions concerning the state's 'organic existence'.

Perhaps surprisingly, Buchanan believes this amounts to an outright refutation of the implications of the social choice impossibility theorems. In particular,

distinction in 'The Pure Theory of Government Finance: A Suggested Approach'.

⁷ibid., p. 496.

⁸ibid., p. 498.

⁹Buchanan, 'Social Choice, Democracy, and Free Markets', p. 116.

¹⁰ibid., p. 117.

¹¹ibid., p. 116.

¹²This argument line of reasoning is pursued more fully in Carol Rovane, *The Bounds of Agency: An Essay in Revisionary Metaphysics* (Princeton: Princeton University Press, 1998), pp. 137–160, and 'What is an Agent?', *Synthese* 140, 1–2 (May 2004): 181–198. Ultimately Rovane rejects such an argument, though for radically different reasons than those I present.

Buchanan claims that Kenneth Arrow succumbs to the metaphysical hokum of 'collective rationality' by positing the social rationality ($\mathcal{SR}$) condition for assessing state decisions by the default rationality thesis. Since Buchanan maintains that the state is not itself a rational agent, he argues that norms of rationality, like those in $\mathcal{SR}$, do not apply to them. So Buchanan concludes that the impossibility results do not impugn democratic processes. In response to these charges, Arrow agrees that it makes no sense to treat the state as an organic entity. Instead, Arrow maintains that 'a rule for arriving at social decisions may be agreed upon . . . without its outcomes being treated as evaluations by anyone in particular'.¹³ Indeed, Arrow agrees that social choices do reduce to the choices of the state's constitutive members—the liberal state's choices come from pooling these individual choices into an all-things-considered judgment—and so Arrow sees himself in accord with Buchanan's philosophy of individualism. There are no 'state' decisions per se; there is only amalgamations of individual decisions, and $\mathcal{SR}$ simply assesses the latter. To quote Isaac Levi, 'Arrow seems to think that in social choice we have choice without a choosing subject and preferences without a preferring subject just as, for Popper, in science we have knowledge without a knowing subject'.¹⁴ So while Arrow believes the applicability of his theorem remains, he does agree with Buchanan that the state should not be regarded as a rational agent.¹⁵

Buchanan and Arrow therefore seem agreed in their denial of the symmetry thesis, suggesting that assessing state decisions by the norms of rationality is both unnecessary and metaphysically suspect. I believe this position has been extremely influential in political science and economics, where

The proper approach to social welfare functions appears to begin with the frank admission that such functions are social, not individual, and therefore are of a fundamentally different philosophical dimension from individual values or from individualistically oriented decision-making processes.¹⁶

Consequently, both disciplines adopt a reductive attitude by understanding collect-

¹³Kenneth J. Arrow, *Social Choice and Individual Values*, 2nd ed. (New York: Wiley, 1963), p. 106.

¹⁴Isaac Levi, *Hard Choices: Decision Making Under Unresolved Conflict* (Cambridge: Cambridge University Press, 1986), p. 152. Levi makes this comparison with Popper because Arrow, *Social Choice and Individual Values*, p. 106, affirms Popper's position on this issue.

¹⁵In fact, theorem 3 in Amartya K. Sen, 'Internal Consistency of Choice', *Econometrica* 61, 3 (May 1993): 495–521, p. 511, appears to demonstrate that $\mathcal{SR}$ is *not* even necessary to prove an impossibility result. However, dropping $\mathcal{SR}$ comes at the expense of strengthening the conditions of non-dictatorship ($\mathcal{D}$) and independence of irrelevant alternatives ($\mathcal{IIA}$) in ways that may be controversial.

¹⁶Buchanan, 'Social Choice, Democracy, and Free Markets', p. 118.

ive action in terms of individual actions. Even so, having laid these two objections out, I believe they are ultimately unpersuasive.

To begin with, these objections overlook the obvious: it is not unusual to hold the decisions of social groups to the same standards as those of individual persons. The purpose of most organizations, and certainly that of the liberal state, is to make and act upon decisions for promoting the values and commitments that its constitutive members deem important. When such a group chooses or acts in a way that would be inconsistent for an individual person to choose or act, there is nothing peculiar or philosophically problematic about criticizing it on those grounds. One is usually not perceived as spouting hokum when charging a group with making an irrational decision. People pass these judgments all the time—especially concerning the government. Hence, it is already common practice to judge the rationality or irrationality of certain groups. In fact, it remains common practice in political science and economics to treat certain groups—families, corporations, and government agencies—as rational agents by holding them to standards of rationality. This alone is very suggestive that it not nonsensical to assess the decisions of groups in a similar manner as those of individual persons.

Nevertheless, a political scientist or economist might still follow Buchanan by maintaining that there is no real need to directly assess the decisions of the state because explaining the state's values, commitments, and resultant choices really just involves explaining all the state's constitutive members' values, commitments, and choices. Everyday talk may sound like it is treating certain groups like rational agents, but this is actually shorthand against having to always break discussion down to the individual members of the group. In principle, it remains sufficient to only discuss individual persons, and only apply norms of rationality to their choices. Therefore, so this rebuttal concludes, political science and economics may rightfully regard talk of the rationality of decisions by the state as unnecessary and ultimately pernicious.

As Isaac Levi rightly argues, however, this reductive attitude easily extends to individual decision making as well.¹⁷ A person is the sum of complex systems that involve a myriad of neurophysiological processes. Nevertheless, I know of few political scientists and economists arguing that it is therefore unnecessary to apply norms of rationality to a person. Reducing an individual person's values, commitments, and choices to biological processes is not required—nor even meaningful—when judging the rationality of a healthy person.¹⁸ Just as the pos-

¹⁷Levi, *Hard Choices*, pp. 151–152.

¹⁸One notable rejection of this view, however, is seen in George Ainslie, *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person* (Cambridge: Cambridge University Press, 1992), and *Breakdown of Will*, (Cambridge: Cambridge University Press, 2001). Ainslie attempts to apply economic principles to the various systems within the individual person,

sible reduction of a person's decision-making to biology need not detract from assessing that person according to norms of rationality, so the possibility that social decision-making may be reduced to that of individuals does not in itself demand excepting the state from these norms as well. This consideration reveals that Buchanan's allegations that it is unnecessary and hokum to treat the state as a rational agent may also apply to individual persons. When talking about individuals, however, these allegations carry little weight. On reflection, then, it is no longer clear why they must always apply to the state.

Even so, there does remain the crucial difference between the individual and the state: the latter has a clear unity of body and consciousness not possessed by the former. All those neurophysiological processes within a person are at least contained within one body and constrained by one consciousness. I do not believe that I need to resort to any controversial notions of metaphysics or psychology to explain why I believe that this distinction is not relevant to what counts as a rational agent. As I understand it, a rational agent is something that possesses values and commitments, which its decisions attempt to uphold and advance. It is this goal of adhering to and promoting these values and commitments that make it sensible to evaluate these decisions with norms of rationality. I see no reason why unity of body or consciousness is required. It is the values and commitments, along with the desire to act in accord with them, that give unity to a rational agent. A person may individually set up and pursue this goal, or a group may collectively attempt to do so. To illustrate, a competent doctor is, after all, committed to, among other things, the health of her patients and the integrity of her practice, and she chooses her options accordingly. The National Institutes of Health (NIH) has similar concerns, only on a much broader scale: the health of Americans and the integrity of the entire profession. Doctor and NIH are each a rational agent in the sense I propose here, and the decisions of either may be judged without reduction to their respective individual parts.

I believe that this reveals that there is some justification for Buchanan's suspicion that Arrow treats the state as a rational agent. Despite the assuring rhetoric of social choices without a choosing subject, Arrow implicitly recognizes that the state actually *does* make choices because it attempts to act upon an all-things-considered weak ordering of the options that is generated by pooling the individual judgments of its constitutive members. As a result, Arrow treats the state as a rational agent because he claims that the state should be committed to making decisions that somehow respect and promote the views of its constitutive members, and do so in accord with the norms of the default rationality thesis. Charles Plott therefore

hence the name he coins, 'picoeconomics'. In particular, he believes that he can use these principles to explain phenomenon like weakness of will.

observes that ‘operationally, it is difficult to distinguish efforts which are motivated from [Arrow’s] point of view from those having the motivation outlined in the [organic point of view]’.¹⁹ Without specifying this additional operational content, speaking of choice without a chooser contributes nothing but metaphysical confusion to the conversation. When a system uses values and commitments to make decisions (as in Arrow’s chooserless choices), and it is appropriate to judge these decisions based on norms of rationality (again as Arrow assumes), then this system is being treated as a rational agent in the sense I am proposing here.

As a result, recognizing the state as a rational agent need not comprise a thick organic conception of the state, nor should it lead to either psychological concerns (for individuals) or mysterious metaphysics (for the state). It is not necessary to suppose a group of deliberating homunculi inside individuals, nor to presume firm physicality or other ontological properties for the state.²⁰ I am only suggesting that individual persons and the state often attempt to make decisions in accord with a set of values and commitments they each deem important. As Levi quite aptly suggests, it is unclear why ‘differences in the “hardware” between the individual and the state entail that only the decisions of the former may be judged by standards of rationality.’²¹ This places the onus is firmly on those who claim that these differences are relevant.

4.2 The Symmetry Between the Individual and the Liberal State

In light of the considerations of the previous section, I see nothing unnecessary or meaningless in assessing state decisions by norms of rationality. I believe there at least as much metaphysical hokum in denying the existence of the obvious as in positing the existence of the fantastic. In any case, it should now be clear why rationally assessing social choice need not rest on any sort of problematic metaphysical views of the state. This is, of course, only the first step in defending the symmetry thesis. So far I have focused on showing that there is nothing metaphysically problematic with treating both individual persons and the liberal

¹⁹Charles R. Plott, ‘Path Independence, Rationality, and Social Choice’, *Econometrica* 41, 6 (November 1973): 1075–1091, p. 1078. Plott, incidentally, does not reject the symmetry thesis: ‘The “degree” of rationality exhibited by [social] choice would, presumably, be the same as that of other “independent” agents, such as individuals, unless, of course, one sees that fellow named Society as having a “deviant personality”’, *ibid.*, p. 1078. Also see Levi, *Hard Choices*, pp. 152–153.

²⁰However, see Rovane, ‘What is an Agent?’, where there is more willingness to take seriously the idea that a single human body may be home to multiple individuals, as in the case of dissociative identity (i.e., multiple personality) disorder.

²¹Levi, *Hard Choices*, p. 152.

state as rational agents. The question remains whether the decisions of both types of agents should be judged by the same norms of rationality. This sets the stage for a further challenge to the symmetry thesis: individual persons, on the one hand, and the liberal state, on the other, possess fundamentally different value and commitments, and this suggests that the same norms of rationality should not apply to both.

In his list of ‘fundamental philosophical issues’ raised by using the norms of the default rationality thesis to assess state decisions, James Buchanan suggests that doing so is simply illogical. This is because he believes that the decisions of the liberal state have a constraint not applicable to those of individuals. Namely, the liberal state must be responsively impartial between the reasonable views of its constitutive members; an individual person is supposed to be responsively *partial* to her own views. Put differently, responsive impartiality requires the liberal state to make decisions in accord with a set of several value orderings, one for each of its constitutive members; the individual person only decides in accord with a single value ordering. Since the default rationality thesis entails that ‘it is legitimate to test the rationality or irrationality of [a social] entity only against [its own, singular] value ordering’,²² this creates problems for the liberal state, which is supposed to be tested against its adherence to a plurality of orderings. From this, he concludes that it is logically inconsistent to judge the decisions of the liberal state by the norms of the default rationality thesis. As Buchanan explains:

If the social group is considered, questions may be raised relative to the wisdom or unwisdom of this organic being. But does not the very attempt to examine such rationality in terms of individual values introduce logical inconstancy at the outset? Can the rationality of the social organism be evaluated in accordance with any value ordering other than its own?²³

Given his proposed asymmetry of purposes between the individual person and the liberal state, he maintains that only the decisions of individuals may be judged by the norms of the default rationality thesis. These norms are not applicable to the liberal state.

Unfortunately, Buchanan provides no formal argument demonstrating a contradiction here, and so it is unclear what in logic prohibits holding both individual persons and the state to the default rationality thesis while also requiring the state to remain responsively impartial. It might be tempting to suggest that Buchanan appeal to the impossibility results of social choice theory to formally back his

²²Buchanan, ‘Individual Choice in Voting and the Market’, p. 116.

²³*ibid.*, p. 116.

claim, but these theorems do not necessitate giving up the symmetry thesis at all. In fact, these theorems appear to establish that holding the symmetry thesis is perfectly logical, though coming at a price. That is, all that the theorems prove is that while individual judgments may satisfy the default rationality thesis (which the individual rationality ($\mathcal{IR}$) condition embodies), if the state remains responsively impartial (which the conditions of weak Pareto ($\mathcal{P}$), non-dictatorship ($\mathcal{D}$), unrestricted domain ($\mathcal{U}$), and independence of irrelevant alternatives ($\mathcal{IIA}$) embody), then state decisions cannot always satisfy the default rationality thesis (at times state decisions will violate the social rationality ($\mathcal{SR}$) condition). In sum, the implication of these theorems is that it is perfectly consistent to hold both the symmetry and the default rationality theses, but conclude that the liberal state is often hopelessly irrational. As it stands, this is not a promising conclusion for defenders of liberal democracy (and one that I reject in section 4.3), but at least it shows there is no logical inconsistency with holding the symmetry thesis.

Despite the fact that there is nothing illogical about holding the symmetry thesis, Buchanan's point that the liberal state has a constraint of responsive impartiality that is not applicable to the individual lingers may appear to be a strong enough reason for the liberal democrat to abandon the symmetry thesis. One need not go as extreme as Buchanan, though, in rejecting the very possibility of assessing the rationality of state decisions. Instead, it might be tempting to simply require that the liberal state adhere to different norms of rationality than the individual person.²⁴

However, I suggest that this is too hasty a conclusion. Indeed, I posit that fairly routine cases of individual decision making have an analogous structure to those of the liberal state. If it is possible for individuals to make rational decisions in these circumstances, and I believe it certainly is, then I see no reason why it must be impossible for the state. To get there, I first explain how there are at least two classes of decision making that require an individual person to adopt an attitude of responsive impartiality in the face of a multiplicity of values and commitments. If both individuals and the liberal state face this same sort of constraint, then Buchanan's concerns no longer tarnish the veracity of the symmetry thesis.

The first and most obvious way in which individual persons are responsively impartial involves situations of fiduciary responsibilities. Consider that individuals, especially those in positions of authority—a president, a senator, a doctor, a parent—must often deliberate and make decisions affecting other individuals. Such a decision maker is not always supposed to employ her own value ordering

²⁴For instance, some game theorists, e.g., Jon Elster, 'The Case for Methodological Individualism', *Theory and Society* 11, 4 (July 1982): 453–482, and some statisticians, e.g., Leonard J. Savage, *The Foundations of Statistics*, 2nd ed. (New York: Dover, 1972), adopt an asymmetrical approach like this.

	Dewey	Cheetham	Howe
1 st	T_1	T_2	T_3
2 nd	T_2	T_3	T_1
3 rd	T_3	T_1	T_2

FIGURE 4.1: Dr. Bailey is committed to the interests of Dewey, Cheetham, and Howe, three patients under her care. She is choosing between treatment allocations T_1 , T_2 , and T_3 , each of which affects these three of the patients in different ways. Above, each allocation is ranked by each patient, according to how Dr. Bailey expects it to benefit that patient.

when deciding; rather she is often expected to serve the values and commitments of these other people. Hence, she must decide by applying a set of orderings comprised of the other individuals' valuations. For instance, consider Dr. Bailey, who has a commitment to Dewey's medical needs along with similar commitments to Cheetham and Howe. As illustrated in figure 4.1, Dr. Bailey has a separate ordering concerning the possible treatment allocations for *each* patient. Unfortunately, this situation involves scarce medical resources, which causes these three orderings to conflict. Satisfying one patient comes at the expense of another, and so Dr. Bailey faces the dilemma of allocating treatments while respecting her commitment to each patient. In many cases, CEOs and CFOs of corporations are in similar positions, as are elected governmental representatives and the bureaucrats they appoint. So it is not just the liberal state that has a commitment to responsive impartiality. Individuals in these situations often explicitly share this commitment.

The second class of decision making requires more discussion. It involves instances of value pluralism, where the person has a collection of values and commitments (not necessarily from fiduciary responsibilities) that she is obliged to consider when making decisions. Consequently, this pluralism could lead the individual to also employ a set of value orderings when making decisions. While technically this is not the exactly same constraint that responsive impartiality places on the liberal state, it is structurally similar because both entail that there is a set of orderings, rather than a single one, to which a decision should be responsive. To illustrate this, suppose that Dr. Webber has several values and commitments grounded in his role as a physician that he seeks his decisions to satisfy. In particular, he cares about beneficence, non-malevolence, and respect for patient autonomy. Sometimes these all rank options the same, but sometimes they conflict, as shown in figure 4.2.²⁵ In this case, he is treating only a single patient and yet for any treatment he chooses, two of his professional commitments agree in prescribing

²⁵The similarities between the profiles of value orderings in figure 4.1 and figure 4.2 is intentional. Both profiles are functionally identical to the one seen in figure 3.5 illustrating Condorcet's paradox.

	Beneficence	Non-Malevolence	Respect for Autonomy
1 st	T_1	T_2	T_3
2 nd	T_2	T_3	T_1
3 rd	T_3	T_1	T_2

FIGURE 4.2: *Dr. Webber possesses commitments to beneficence, non-malevolence, and respect for patient autonomy, and he is choosing between treatments T_1 , T_2 , and T_3 for a single patient. The treatments are ranked by each commitment as depicted above.*

another one. So if, for example, he selects treatment T_1 , his commitment to beneficence is upheld but both non-malevolence and respect for autonomy agree that T_3 is a superior choice. Consequently, Dr. Webber's professional commitments generate a troublesome conflict. So just like the liberal state, Dr. Webber is seeking to make a decision that is responsive to a set of value judgments and not a single one. Such decisions are quite common, occurring in even the most mundane of cases, such as buying a house, selecting an insurance policy, or any other instance where the potential actions have multiple attributes that are salient to the decision maker.

Often, both situations may overlap, as in the case of the competent doctor who has both fiduciary commitments to the health of her patients, as well as a plurality of other values related to the integrity of her practice. In this manner, there does not seem a significant difference between that doctor and an organization like the National Institutes of Health. Both attempt to make decisions that are responsive to a whole host of competing values and commitment. The fact that one is individual and the other is composite does not seem relevant, for both appear committed to responsive impartiality. As a result, there is no need to claim that different norms of rationality should apply for each.

Even so, this appeal to the symmetry thesis, especially in the instance of value pluralism, might seem to fall short. After all, I presented (in section 1.2) the pervasive idea that a reasonable doctrine ought to be capable of resolving potential conflicts between its various values and commitments. Systematicity, in other words, demands that Dr. Webber, or anyone else in such a situation, revise his doctrine so that it determines the proper weighing and ordering of these competing commitments in order to arrive at a single, all-things-considered judgment. For if he does not do this, so it is suggested, Dr. Webber cannot avoid choosing a treatment in an ad hoc or arbitrary manner. Similarly, it might be thought that Dr. Bailey must strive for systematicity in her assessments of the three patients so that she can make a principled decision. In general, a systematic doctrine attempts to reduce a plurality of value orderings into a single one, which in turn allows for

a decision to made in accord with the default rationality thesis.²⁶ An unsystematic doctrine is an embarrassment requiring repair. And while is it perfectly acceptable for a person to make their views more systematic, it is unacceptable for the liberal state to do so because, as I have already said (in section 1.4), doing so comes at the expense of responsive impartiality. As a result, the concern remains that in instances of value pluralism, at least, individual persons are fundamentally different from the liberal state, and so the symmetry thesis does not apply to them.

My primary concern with this argument against the symmetry thesis is that it does not appreciate that an individual, just like the liberal state, must often make decisions in the face of unresolved conflicts. For instance, Dr. Webber's different commitments might not easily commensurate into a single ordering for ranking the treatments. This pluralism might take two different forms. It could be an *open* pluralism, where Dr. Webber is epistemologically incapable of determining the appropriate way to weigh and order his commitments to beneficence, non-malevolence, and respecting patient autonomy. Further reflection and inquiry is necessary for Dr. Webber to set these weights, but, unfortunately, there is no time for doing any of this. A choice must be made now. Or it may be a *closed* pluralism, where Dr. Webber is, as a matter of principle, unwilling to set a precise weighing or ordering between these commitments. Perhaps he maintains that these commitments are so radically different, and that trying to determine weights would be as fruitless as 'comparing apples and oranges'.²⁷ In either case, Dr. Webber is unable or unwilling to commensurate beneficence with respecting patient autonomy into a single all-things-considered value ordering.²⁸ In a similar manner, Dr. Bailey could be unable or unwilling to commensurate her commitments to each of her patients.

In such instances of unresolved conflict, I posit that something like responsive impartiality is necessary when making a decision. This is probably most clear for Dr. Bailey, who wants to makes decisions that are not only responsive to each of her patients' needs but also impartial between her patients. Favoring one patient over the other would certainly resolve the conflict, but at the obvious expense of impartiality. Similar reasoning extends to Dr. Webber's unresolved conflict.

²⁶Later on (in section 4.3), I refer to this idea that a person should systematically make judgments for transforming a plurality of value orderings into a single, all-things-considered ordering, as the extended rationality thesis.

²⁷A very similar instance of closed pluralism for clinical research appears in Benjamin Freedman, 'Equipose and the Ethics of Clinical Research', *New England Journal of Medicine* 317, 3 (16th July 1987): 141–145, p. 143, where he argues against what he calls 'theoretical equipose' on the grounds that it 'tantamount to the formulation of a rigorous calculus of apples and oranges'.

²⁸A similar distinction, between 'tentative' and 'assertive' incompleteness, appears in Amartya K. Sen, 'Maximization and the Act of Choice', *Econometrica* 65, 4 (July 1997): 745–779, pp. 763–764, and 'Incompleteness and Reasoned Choice', *Synthese* 140, 1–2 (May 2004): 43–59, p. 55.

Not only should he be responsive to each of his conflicting commitments, but he should also be impartial between them. Showing partiality towards, say, respecting patient autonomy does resolve the conflict, but this is something which Dr. Webber cannot or will not do at the moment of choice. In both cases, an individual is committed to what amounts to responsive impartiality, much like the liberal state when it is navigating the inevitable unresolved conflicts between the reasonable views of its constitutive members. Consequently, the symmetry thesis obtains.

The underlying case against the symmetry thesis is that individuals should be partial while the liberal state should be impartial. However, in cases of unresolved conflict, I believe that both should remain responsively impartial. The presumed partiality of individuals rests, I believe, on solely focusing on instances where the individual is not so conflicted. For instance, had Dr. Bailey's patients all ranked the treatment allocations in the same order, then there would be no need for impartiality because favoring one patient means favoring all. Similarly, had all of Dr. Webber's commitments lined up in agreement with their valuations, there would be no need to be impartial between them. In these cases of unanimity amongst commitments, I agree that there is no need to be impartial. Notice, however, that the same is true for the liberal state: when there is no social controversy, the state need not, strictly speaking, be impartial. Responsiveness, by acting on the unanimous judgment, is sufficient. Once again, the symmetry thesis remains viable.

4.3 Impossibility Theorems and the Symmetry Thesis

Having defended the symmetry thesis, I believe most of the pieces are now on the table for presenting my argument against the pessimistic implications of the impossibility theorems of social choice theory for liberal democracy. Ultimately, these implications rely on the assumption that the norms of the default rationality thesis embody the correct way to assess the decisions of a rational agent. The symmetry thesis, however, suggests that this notion of rationality has substantial flaws when it comes to decision making under unresolved conflict. In this section, I flesh out what this all means and why it stands against the impossibility results.

While James Buchanan may also protest the pessimistic implications of the impossibility results, there remain some significant commonalities between Buchanan and these results worth emphasizing: they all (1) identify a problem with applying the norms of the default rationality thesis to the decisions of the liberal state, (2) presume there is no such problem for individual persons, and (3) suggest that this is because the liberal state has a constraint of responsive impartiality not applicable to the individual. From here, they diverge, with Buchanan denying that the

liberal state is a rational agent and the theorems implying that the liberal state is often unavoidably irrational when attempting to pool individual judgments into a collective, all-things-considered assessment. Contrary to (3), I maintain that responsive impartiality is applicable to both individual persons and the liberal state. Consequently, whatever is a problem for the liberal state in making rational decisions must also be a problem for the individual as well, and whatever is not a problem for individual decision making should not be a problem for the liberal state.

This last point is important, for it turns out that the implications of Kenneth Arrow's theorem may be more pessimistic and far reaching than even he may recognize.²⁹ As my examples of the doctors from the previous section show, individuals in situations of unresolved conflict often must make decisions structurally similar to those of the liberal state. These doctors are attempting to remain responsively impartial to a collection of value orderings, instead of acting partial to a single one. This evidentially implies that Arrow's theorem applies to their respective decisions as well.

For instance, the impossibility results would appear to suggest that Dr. Bailey from figure 4.1 cannot rationally choose an optimal treatment in accord with the default rationality thesis that is responsive to her commitments to each patient. That is, if Dr. Bailey seeks to promote each of her patients' interests, she should recognize, within limits, any possible combination of weak orderings for the patients (satisfying the individual rationality ($\mathcal{IR}$) and unrestricted domain ($\mathcal{U}$) conditions), and then always make her decision relative to a single, all-things-considered, weak ordering the options (satisfying the social rationality ($\mathcal{SR}$) condition, though in this case, the all-things-considered ordering is not 'social' in the usual sense). Furthermore, her ordering should only be based on how each patient orders the actual options and not how they are ranked relative to the merely possible ones (satisfying the independence of irrelevant alternatives ($\mathcal{IIA}$) condition) and honor any consensus among the patients (satisfying the weak Pareto ($\mathcal{P}$) condition). Finally, her decisions should not always display partiality to one of the patient's interests over

²⁹In this and the next two paragraphs, I show how Arrow's theorem applies to individual decisions. Similar examples are easily constructed to show how the theorem from Teddy Seidenfeld, Joseph Kadane, and Mark Schervish also applies to individuals as well. For instance, see Teddy Seidenfeld, 'Outline of a Theory of Partially Ordered Preferences', *Philosophical Topics* 21, 1 (Spring 1993): 173–189; Teddy Seidenfeld, Mark J. Schervish and Joseph B. Kadane, 'A Representation of Partially Ordered Preferences', *Annals of Statistics* 23, 6 (December 1995): 2168–2217, and 'Coherent Choice Functions Under Uncertainty', *Synthese* 172, 1 (January 2010): 157–176; and Mark J. Schervish, Teddy Seidenfeld, Joseph B. Kadane and Isaac Levi, 'Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons', in *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, ed. by Jean-Marc Bernard, Teddy Seidenfeld and Marco Zaffalon (Carleton Scientific, 2003): 496–510.

the others' (satisfying the non-dictatorship ($\mathcal{D}$) condition). If these conditions hold, then so does Arrow's theorem. Thus, it is impossible for Dr. Bailey to be both responsively impartial and make decisions consistently satisfying the norms of the default rationality thesis.

Recall that I noted (in section 3.2) that the formal framework of Arrow's theorem need not exclusively apply to social choices, though that is usually where the focus of the theorem is usually understood. Rather, the apparatus is general enough that it applies to *all* instances of pooling a set of value judgments into a single, all-things-considered judgment. So the pessimistic implications of the social choice theorems are not limited to social choice, but bleed into the domain of value pluralism. So Dr. Webber from figure 4.2 is also in a dilemma because of his various conflicting commitments, even though he is only dealing with one patient. Once again, Arrow's conditions appear as reasonable constraints on how Dr. Webber should choose, given that Dr. Webber wants to remain responsive to all these obligations. So he should be prepared to take any possible combination of weak orderings dictated by his commitments (satisfying $\mathcal{IR}$ and $\mathcal{U}$) and arrive at a single, all-things-considered ordering of the options (satisfying $\mathcal{SR}$) based only on how his commitments order the actual options (satisfying $\mathcal{JJA}$), while honoring any consensuses among these commitments (satisfying $\mathcal{P}$) and not allowing the same commitment to always dominate (satisfying $\mathcal{D}$). Once again, if this is the case, then the theorem applies, and so it is impossible for Dr. Webber to consistently make rational decisions because he cannot always derive a single, all-things-considered ordering of the treatments.³⁰ In situations of unresolved conflict between their various values and commitments, all individuals apparently suffer from this affliction.

In these respects, the seemingly pernicious implications for the liberal state applies equally to individual decision making under unresolved conflict. Both the individual and liberal state will always face situations where, according to Arrow's theorem, it is logically impossible to make a responsively impartial decisions in accord with the default rationality thesis. Of course, most theorists, including Arrow and Buchanan, do not then conclude that these individuals are unavoidably irrational. In fact, Arrow's assumption of 'non-triviality' presumes that it is only when there is more than one person involved that impossibility looms. But if the impossibility results of social choice still apply to groups, and not individuals, it seems that individuals must do something special in order to make rational decisions that the liberal state cannot. For reasons similar to what I suggested

³⁰Susan L. Hurley, *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989), pp. 226–241, recognizes a similar idea of applying the social choice conditions to instances of individual choice under conflict.

earlier (in section 4.2) regarding doubts concerning the symmetry thesis, I believe the assumption here is that the individual can reason more systematically while reasonable pluralism prevents the liberal state from doing so. In particular, I suspect that these theorists assume that an individual possesses additional information that governs how her various values and commitments interact, which in turn allows the individual to adopt a more systematic view in response to internal conflict.³¹

The major source of information underlying many types of interactions between values and commitments involves intrapersonal comparisons of value. That is, an individual may not only compare the options according to how each of her commitments ranks them but also compare these options across these commitments. Returning to the situation of Dr. Bailey, if Arrow's theorem applies, there is no way for her to choose which treatment allocation to administer while remaining responsively impartial between her commitment to her patients. However, additional information may be readily available for her to help her form a single, all-things-considered assessment of the options. For instance, she may judge that treatment allocation T_1 would be more beneficial to Dewey than allocation T_2 would be for Cheetham given such factors as the patients' respective ages and medical histories. The particular content of Dr. Bailey's intrapersonal comparisons would, of course, depend on additional information associated with the allocations, how they are administered, features of the patients, and so forth. Intrapersonal information might also be available for Dr. Webber as well: he may judge that treatment T_1 with respect to beneficence is superior to treatment T_2 with respect to non-maleficence because the lesser risks of debilitating side-effects from administering T_2 are outweighed by T_1 's especially beneficial properties. Again, the particular features of the situation would determine the precise content of these evaluations for Dr. Webber.

In making intrapersonal evaluations, an ordering is generated, though not one over the options per se. Rather, the ordering is over the ordered pairs of coupling one option with one commitment. So, for instance, Dr. Bailey may rank T_1 for Dewey (one ordered pair) superior to T_2 for Cheetham (the second ordered pair). Similarly, Dr. Webber may rank T_1 for beneficence (one ordered pair) above T_2 for non-maleficence (the second ordered pair). Furthermore, just like a regular value ordering over options alone, an ordering generated by intrapersonal comparisons may also comprise a ranking that either only contains ordinal information or also contains cardinal, that is, numerical, information as well (recall this distinction from section 3.1). In general, then, it is sometimes possible for an individual to

³¹Robert Nozick, 'Moral Complications and Moral Structures', *Natural Law Forum* 13 (1968): 1–50 offers an extensive treatment of various ways that values and commitments can interact with each other.

appraise the options across her various commitments in order to ascertain the information necessary for making intrapersonal comparisons.

Once the individual identifies this information, the idea is that she may then use it to systematically arrive at a single, all-things-considered ordering of the options. One common decision-making method that does this is conjoint measurement analysis.³² Using intrapersonal information, this method determines the trade-offs that an individual is willing to make between her values and commitments. This facilitates the sort of weighing and ordering of commitments that systematicity advocates. Hence, given the ages and medical histories of Dewey and Cheetham, Dr. Bailey may be willing to trade-off the benefits to one patient with greater benefits to the other. Similarly, Dr. Webber may be willing to trade a certain level of risks (considering non-maleficence) for a certain level of benefits (considering beneficence). Once these trade-off schedules are determined, then it is possible to arrive at a 'common denominator' for constructing a single, all-things-considered judgment of the relative merit of each options. Intrapersonal information therefore can reduce a set of orderings evaluating the options into just one, with which the individual uses to select the optimal option. Unlike the default rationality thesis, this approach, and others like it that employ interpersonal information, do not presume that rationality consists of optimization of a pre-specified weak ordering. Instead, they posit that rationality involves first systematically deriving that single ordering from a set of orderings, and then optimizing. I call this the *extended* rationality thesis. Nevertheless, as is seen in figure 4.3, this new conception of rationality is an extremely conservative move: the default rationality thesis is just prefaced with an account of using intrapersonal comparisons to reduce several orderings into a single one.

If similar information exists for the liberal state, then it may perform *interpersonal* comparisons of value assessments that satisfy the extended rationality thesis.³³ For instance, consider figure 4.4 where the liberal state is choosing between three different health care policies. This group faces a dilemma symmetrical to those of the doctors: the liberal state is committed to remaining responsively impartial between these coalitions' reasonable judgments, but these judgments conflict with one another. Hence, Arrow's theorem holds that the conflict cannot be simply

³²Ralph L. Keeney and Howard Raiffa, *Decisions with Multiple Objectives: Preference and Value Tradeoffs* (Cambridge: Cambridge University Press, 1993), is the canonical text on this issue, though Kenneth J. Arrow and Herve Raynaud, *Social Choice and Multicriterion Decision-Making* (Cambridge, MA: MIT Press, 1986), also deals with it as well.

³³Allan Gibbard, 'Interpersonal Comparisons: Preference, Good, and the Intrinsic Reward of a Life', in *Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 165–193 also explores the relationship between intra- and interpersonal comparisons, though in a substantially different way than I do here.

Default Rationality Thesis	Extended Rationality Thesis
A rational agent ought to:	A rational agent ought to:
(1) Possess a weak (or cardinal) ordering, and then	(1) Possess a set of weak (or cardinal) orderings,
(2) Select its optimal element(s).	(2) Derive a weak (or cardinal) ordering from this, and then
	(3) Select its optimal element(s).

FIGURE 4.3: *The default rationality thesis compared with the extended rationality thesis.*

	Coalition 1	Coalition 2	Coalition 3
1 st	P_1	P_2	P_3
2 nd	P_2	P_3	P_1
3 rd	P_3	P_1	P_2

FIGURE 4.4: *The liberal state is choosing to implement a health care policy. Given the choice between P_1 , P_2 , and P_3 , the above profile of value judgments represents the views of three equally sized coalitions that divide the liberal state.*

resolved by a pooling function. Despite this, the liberal state must still make a decision. However, in a similar manner as each of the doctors, the liberal state might possess uncontroversial information that allows it to compare how the members of each coalition fare relative to the members of the other coalitions. That is, the state might ascertain that coalition 1 fares better under plan P_1 than coalition 2 does under plan P_2 . Perhaps the members of coalition 1 are at higher risk for certain ailments that plan P_1 would cover, whereas the members of coalition 2 have remarkable health and so little need for any health care plan. Once again, the potential information varies with the particulars of the situation.

In an earlier discussion of *JJA* (in section 3.6.1), I noted that this condition filters out such information, but that the condition is easily modified to accommodate it. This modification then generates a possibility theorem for social choice, which shows that methods like summation and maximin can be responsively impartial while also satisfying what I am now calling the extended rationality thesis. But I also argued that this possibility may quickly return to impossibility because of reasonable social controversy within the liberal state concerning the proper metric for interpersonal comparisons, the data measured by any given metric, and the proper way to use this data for making a social choice.

In fact, this very concern also makes an appearance in Buchanan's list of 'fundamental philosophical issues', where he argues that the very concept of a common denominator for the liberal state to compare options is 'vague and general', offering 'little or no direct guidance' for any practical application.³⁴ This is justified by two radically different arguments. The first maintains that interpersonal comparisons are impossible.³⁵ This is justified by claiming that the values and commitments of a person are, at bottom, derived from the subjective preferences of that person. Since the fact of reasonable pluralism recognizes that individuals may possess different value orderings, it is therefore meaningless to compare the satisfaction of one person's value ordering with that of another. As Lionel Robbins explains, 'introspection does not enable A to measure what is going on in B's mind, nor B to measure what is going on in A's' and so 'there is no way of comparing the satisfaction of different people'.³⁶

Recently, though, this view has fallen out of favor by many theorists.³⁷ Many political theorists now believe that there do exist metrics, apart from preference satisfaction, that permit interpersonal comparisons. However, this leads to the second argument against interpersonal comparison, which I presented earlier (again, in section 3.6): there is an embarrassment of riches generating a host of actual, reasonable disagreements. The fact of reasonable pluralism therefore ensures that nothing grounds a meaningful space in which to conduct interpersonal comparisons; the liberal state cannot then satisfy the norms of the extended rationality thesis.

The point I wish to emphasize here is that, once again, most political scientists and economists do not see any problem for the individual to satisfy the extended rationality thesis. The underlying assumption is most likely grounded on the idea that an individual person can more easily decide on what sort of metric to employ for her own deliberations, the data that metric captures, and how to apply that data with a method, like conjoint analysis, for making a decision in accord with the extended rationality thesis. The view keeps arising that that individuals do things that the liberal state cannot. In this case, it is the inability of the liberal state to adopt a systematic doctrine in the face of reasonable pluralism that implies that

³⁴Buchanan, 'The Pure Theory of Government Finance: A Suggested Approach', p. 498.

³⁵For instance, see Lionel Robbins, *An Essay on the Nature and Significance of Economic Science*, 2nd ed. (London: Macmillan, 1935), and John R. Hicks, 'The Foundations of Welfare Economics', *Economic Journal* 49, 196 (December 1939): 696–712. Arrow, *Social Choice and Individual Values*, also appears to accept this as well.

³⁶Robbins, *An Essay on the Nature and Significance of Economic Science*, p. 140.

³⁷Donald Davidson, 'Judging Interpersonal Interests', in *Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 195–211, for one, gives reasons for why preference satisfaction does not necessarily render interpersonal comparisons meaningless.

the latter cannot appeal to interpersonal comparisons in order to make decisions. Therefore, the impossibility theorems resurface and the liberal state's decisions are unavoidably irrational according to either the default or extended rationality thesis.

I believe that my earlier argument (from section 4.2) against the denial of the symmetry thesis applies here. Many political scientists and economists repeatedly fail to fully appreciate that individuals must often make decisions in the face of unresolved conflict. Dr. Bailey might be unable or unwilling to find a common denominator for comparing the interests of her patients. Dr. Webber might be in a similar situation relative to his professional commitments. As a result, these doctor's might claim that they are unable or unwilling to apply the tools of conjoint measurement analysis or any other approach for increasing the systematicity of their respective views. Systematicity and the rationality theses respond that in order to make a rational decision, an individual must first resolve all conflicts that may exist between her various values and commitments in order to arrive at an all-things-considered judgment. According to the standard view in political science and economics, unresolved conflict unavoidably leads to inconsistency and incoherence.

This is where the standard accounts of rationality leave us: the greater the pluralism and unresolved conflict within an individual or the liberal state, the more ad hoc, arbitrary, inconsistent, and incoherent the respective individual and collective decisions will be. The symmetry thesis may abide, but conflicted individuals and democracies are in the same impossible position, having to make decisions without any normative guidance.

As a way of concluding this chapter, I posit turning the implications of the impossibility theorems of social choice on their head. Rather than closing the door to rational decision making under unresolved conflict, these results should be seen as searing indictments of the standard accounts of rationality. Reason ought to provide normative guidance for how an individual ought to choose in all circumstances. It should be embarrassing for reason to abandon a person or the liberal state to unavoidable irrationality precisely when there is the most need for assistance in making a principled decision. For reason has little to do when there are no conflicts between commitments or individuals; the optimal choice is then obvious. When circumstances require a hard choice in the face of an intractable pluralism of commitments, it should therefore count against an account of rationality when it refuses or is unable to provide any norms for making a principled decision.

In the following concluding chapter of this dissertation, I suggest that an account of rational decision making need not leave either the individual or the liberal state in such dire straights.

CHAPTER FIVE

Agreement and Legitimacy

RATIONAL decision making need not always require optimization of a single weak ordering as suggested by systematicity and required by the default rationality thesis. Often individuals must make decisions in situations of unresolved conflict, where they have divergent yet incommensurable values and commitments ranking the available options. That is, a person often finds that she must choose amongst options when (1) her various values and commitments conflict and (2) she is either unable or unwilling to reduce these values and commitments to a single, all things considered ranking of her options. Systematicity seems embarrassed by pluralism without priorities or weights, and the default rationality thesis presumes that making a decision in such circumstances is unavoidably arational. In the last chapter I argued that such a claim is surely too strong: a theory of rationality ought to account for how a person may systematically reason when conflicted. Such a theory should explain rational choice based on *multiple* value orderings and not simply on a single one.

I also maintained in the last chapter that the same norms of rationality, whatever they are, ought to apply to both individual decisions and those made by certain groups of individuals. So I maintain that if an account of rational choice under unresolved conflict applies for an individual, it ought to apply to groups in similar circumstances. In fact, proposing that the liberal democratic state make coherent decisions while simultaneously remaining neutral between the various competing comprehensive doctrines its members endorse appears to presume the possibility of a theory of rational choice under unresolved conflict. For like individuals, the liberal democratic state must often make decisions based on conflicting value orderings. As I have presented them, the impossibility theorems demonstrate that

such decisions face difficulties when the state must first resolve these conflicts by reducing any competing values and commitments to a single ranking and then decide based on this, as required by the default rationality thesis.

The issue remains, however, to clarify norms of rational choice under unresolved conflict. Having done that, these norms can then be adapted and applied to the decisions of liberal democratic institutions. Therefore, in this final chapter I take some first steps in this direction by suggesting how rival value orderings may contain sufficient agreements for eliminating some of the available options. By providing a procedure for eliminating options, a decision may become more tractable, though in cases of severe conflict there is only so much a procedure can guarantee. In particular, I argue that an option should be rejected if it is not optimal according to at least one value ordering representing a permissible resolution or compromise to the conflict. Such a consensus of non-optimality should be respected. Furthermore, I maintain that permissible resolutions and compromises must themselves respect certain agreements, when they hold, between the original, conflicting value orderings. Foremost among these is the weak Pareto principle ($\mathcal{P}$) from section 3.2: if the value orderings agree that one option is superior to a second, any compromise should reflect this. Towards the end of section 5.2, I present two additional examples of agreements that permissible resolution should also respect. In the end, though, these sorts of agreements will be extremely context-sensitive and purely procedural considerations may not be able to anticipate them all in advance.

My position is similar to those suggested by Amartya Sen's categorical maximality and Isaac Levi's admissibility. I reject both, however. Categorical maximality does not take into account interpersonal or cardinal information of the sort I discussed in the last chapter, and this makes it too limited in the sorts of resolutions to a conflict it may consider. Admissibility as understood by Levi, on the other hand, does take advantage of this sort of information when it is available, but it ends up being too permissive in the sorts of resolutions it considers. I present these views in sections 5.1 and 5.2 respectively. In order to position myself more clearly between these two views, I argue in section 5.3 that in cases of unresolved conflict a decision-making procedure ought to focus on agreements within the original, conflicting orderings themselves and need not also be responsive to all the potential conflicting reasons underlying these ordering. This is not to say that these reasons are irrelevant—they should certainly be considered and reflected upon during deliberation preceding any decision procedure. But the process for pooling individual judgments need not consider these at the time of choice. Finally, in sections 5.4 and 5.5, I conclude by reflecting on what this means for liberal democratic theory. In sum, the proceduralism for decision making that I endorse does not pretend to resolve social conflict. It only points out the options that

the state may consistently choose in times of unresolved social conflict while remaining neutral between the comprehensive doctrines of its constitutive members. A voting scheme may do this, but the outcome of such a process should not be understood as representing any sort of solution or social consensus to controversial social issues.

5.1 Maximality

In his ongoing analysis of rational decision making, Amartya Sen proposes weakening the default rationality thesis by suggesting that rational choice under unresolved conflict need only involve maximization of a quasi-ordering.¹ Recalling the terminology from section 3.1, a weak ordering sets up a reflexive, transitive, and complete relationship between options. A quasi-ordering, on the other hand, drops the completeness requirement.² By not requiring completeness, quasi-orderings admit the possibility that two or more options may be incomparable. So, for instance, suppose that Jones is an art lover who has a quasi-ordering R_{Jones} concerning artistic merit. Such a value (quasi-)ordering may judge that Wolfgang Amadeus Mozart's *The Marriage of Figaro* and Leonardo da Vinci's *Mona Lisa* are each superior to Antonio Salieri's *The School of Jealousy*.³ Furthermore, it holds that neither *Figaro* nor the *Mona Lisa* is superior to the other, for Jones believes both are truly superb works of art. However, given the radical difference between these two—one is an opera and the other a painting—Jones is not comfortable expressing equivalence between them. As a result, R_{Jones} does not establish any of these relationships between *Figaro* and the *Mona Lisa*. These two works of art are regarded as incomparable, and so R_{Jones} does not satisfy completeness. For a

¹Amartya K. Sen, 'Quasi-Transitivity, Rational Choice and Collective Decisions', *Review of Economic Studies* 36, 3 (July 1969): 381–393; *Collective Choice and Social Welfare*, (Amsterdam: North-Holland, 1970); 'Interpersonal Aggregation and Partial Comparability', *Econometrica* 38, 3 (May 1970): 393–409; 'Interpersonal Comparison and Partial Comparability: A Correction', *Econometrica* 40, 5 (September 1972): 959; *On Economic Inequality*, (Oxford: Clarendon Press, 1973); *On Ethics and Economics*, (Oxford: Blackwell, 1987); *Inequality Reexamined*, (Oxford: Oxford University Press, 1992); 'Internal Consistency of Choice', *Econometrica* 61, 3 (May 1993): 495–521; 'The Formulation of Rational Choice', *American Economic Review* 84, 2 (May 1994): 385–390; 'Rationality and Social Choice', *American Economic Review* 85, 1 (March 1995): 1–24; 'Maximization and the Act of Choice', *Econometrica* 65, 4 (July 1997): 745–779; and 'Incompleteness and Reasoned Choice', *Synthese* 140, 1–2 (May 2004): 43–59.

²Quasi-orderings are sometimes also known as pre-orderings. Regardless, I should emphasize that all weak orderings are quasi-orderings, but not the other way around. Simply put, a weak ordering is just a connected quasi-ordering.

³Salieri was contemporary of Mozart. His fame today rests not so much on the quality of his compositions, but on the dramatic—and presumably highly fictionalized—account of his antagonistic relationship with Mozart in the play (1979) and film (1984) *Amadeus* by Peter Shaffer.

quasi-ordering, such a possibility is permitted.⁴

The problem with a quasi-ordering, though, is that it may not identify an optimal option for choice. Recall from section 3.1 that an optimal option is expressly better than or equally valuable as every other available option according to the value ordering. In the case of the value ordering R_{Jones} assessing artistic merit, for example, there is no optimal option. *The School of Jealousy* is ruled out because it is judged inferior to either of the other two. *Figaro* is not optimal either because it is not explicitly judged better than or equally valuable as the *Mona Lisa*, while symmetrical reasoning also rules out the *Mona Lisa* as optimal.

Even so, a quasi-ordering R always contains at least one ‘maximal’ element. Once again using the terminology and notation from section 3.1, when selecting from a set of available options A , the maximal options are those options not expressly inferior to any other option in A according to value ordering R . These maximal options are denoted by the set $M(A, R)$, which is defined as follows:

Maximization: $M(A, R) = \{x \mid x \in A \text{ and there is no } y \in A \text{ such that } yPx \text{ holds}\}.$ ⁵

As defined here, maximization is weaker than optimization because a maximal

⁴Willingness to drop completeness appears in a smattering of contexts. For instance, see Robert J. Aumann, ‘Utility Theory without the Completeness Axiom’, *Econometrica* 30, 3 (July 1962): 445–462, and ‘Utility Theory without the Completeness Axiom: A Correction’, *Econometrica* 32, 1–2 (January 1964): 210–212; Joseph Raz, *The Morality of Freedom* (Oxford: Oxford University Press, 1986), pp. 335–340, and ‘Incommensurability and Agency’, in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 110–128; Paul Anand, ‘Are the Preference Axioms Really Rational?’, *Theory and Decision* 23, 2 (September 1987): 189–214; Teddy Seidenfeld, ‘Decision Theory without “Independence” or without “Ordering”: What is the Difference?’, *Economics and Philosophy* 4, 2 (October 1988): 267–290; Elizabeth S. Anderson, *Value in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1993), pp. 44–64, and ‘Practical Reason and Incommensurable Goods’, in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 90–109; and Hilary Putnam, ‘Über die Rationalität von Präferenzen’, *Allgemeine Zeitschrift für Philosophie* 21, 3 (1996): 209–228, translated as ‘On the Rationality of Preferences’ in *The Collapse of the Fact/Value Dichotomy and Other Essays*, (Cambridge, MA: Harvard University Press, 2002), pp. 79–95.

⁵Sen, *Collective Choice and Social Welfare*, p. 9. See also Hans G. Herzberger, ‘Ordinal Preference and Rational Choice’, *Econometrica* 41, 2 (March 1973): 187–237, p. 197, where it is referred to as ‘liberal maximalization’. Many of its formal properties are discussed in Thomas Schwartz, ‘Choice Functions, “Rationality” Conditions, and Variations on the Weak Axiom of Revealed Preference’, *Journal of Economic Theory* 13, 3 (December 1976): 414–427; Kotaro Suzumura, ‘Remarks on the Theory of Collective Choice’, *Economica* 43, 172 (November 1976): 381–390, and *Rational Choice, Collective Decisions, and Social Welfare*, (Cambridge: Cambridge University Press, 1983); Kotaro Suzumura and Yongsheng Xu, ‘Recoverability of Choice Functions and Binary Relations: Some Duality Results’, *Social Choice and Welfare* 21, 1 (August 2003): 21–37; and Mark A. Aizerman, ‘New Problems in the General Choice Theory: Review of a Research Trend’, *Social Choice and Welfare* 2, 4 (December 1985): 235–282.

option need not expressly beat every other option; there is just no other option that beats a maximal option. So returning to value ordering R_{Jones} , *Figaro* and the *Mona Lisa* are both maximal, even though neither is optimal. However, it is worth noting that when the value ordering involved is a weak ordering, then the optimal and maximal options are the same.⁶

So far this account still only involves decision making with respect to a single value ordering. When dealing with a profile of value orderings $\mathcal{R} = \langle R_1, R_2, \dots, R_n \rangle$, though, Sen suggests that it can be reduced to a single ordering R by employing the following rule:

Intersection: xR_1y holds if and only if for each value ordering $R_i \in \mathcal{R}$, xR_iy holds.⁷

If each of the orderings in $\mathcal{R}$ is a weak ordering, then the result will be a quasi-ordering R_1 that respects all unanimous agreements of the form that ‘option x is superior to or equally valuable as option y ’. Maximizing according to the quasi-ordering generated by intersection yields the following rule of choice amongst available options in A given a profile of value orderings $\mathcal{R}$:

Intersection Maximization: $M_1(A, \mathcal{R}) = \{x \mid x \in A \text{ and there is no } y \in A \text{ such that } yP_1x \text{ holds}\}$, where P_1 is from the intersection R_1 of $\mathcal{R}$.⁸

This decision rule eliminates an option when an alternative is found to be better than it according to the intersection of value orderings in $\mathcal{R}$. The idea is that when every value ordering ranks an option higher, this unequivocally shows that this option is better regardless of what priorities or weights these orderings may have. This allows the elimination of options even though there is disagreement over these priorities and weights. To illustrate, consider the value orderings of Barry and Michelle in figure 5.1. The only consensus here is that coal power is superior to nuclear power, and so the intersection only contains that one judgment. Consequently, intersection maximization eliminates nuclear power without requiring Barry and Michelle to determine how to weigh their respective value orderings. Of course, coal power and hydropower remain maximal, so Barry and Michelle may still have a hard choice to make, but at least some progress has been made for that decision.

⁶For the proof, see Theorem A(2)(b) in Suzumura, *Rational Choice, Collective Decisions, and Social Welfare*, p. 13.

⁷This is the rule presented in Sen, *On Economic Inequality*, pp. 72–74, though Sen limits it for profiles $\mathcal{R}$ of total orderings. This is an important consideration, so I return to it in footnote 12 in this chapter.

⁸‘Incompleteness and Reasoned Choice’, pp. 49–50. It is called the ‘intersection approach’ in *On Economic Inequality*, pp. 120–121, and *Inequality Reexamined*, pp. 46–49. A similar approach is implicitly seen in the discussion of the Atkinson measure of inequality in Anthony B. Atkinson, ‘On the Measurement of Inequality’, *Journal of Economic Theory* 2, 3 (September 1970): 244–263, pp. 258–261.

	Barry	Michelle
1 st	coal	hydro
2 nd	nuclear	coal
3 rd	hydro	nuclear

FIGURE 5.1: *Barry and Michelle's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower.*

As is, I worry that intersection maximality involves an unsatisfactory process for eliminating available options. This is because intersection maximality excludes an option provided that (1) all the value orderings agree that it is either inferior to or equally valuable as some alternative in A and (2) at least one ordering judges that option inferior to that alternative.⁹ This shows adherence to what is known as the strong Pareto principle, and its effects are seen in George and Laura's situation in figure 5.2.¹⁰ In that case, only nuclear power is eliminated because (1) both George and Laura agree that it is either inferior to or equally valuable as coal power and (2) George judges that it is inferior to coal power. Neither condition holds for coal power, so it remains intersection maximal. The apparent the idea here is that Laura's indifference between coal power and nuclear power allows George's more affirmative judgment to decide the matter.

My concern here is with the sort of agreement intersection is trying to capture. Intersection begins by preserving a unanimous agreement concerning a disjunction: *either* coal power is superior to nuclear power *or* coal power is equally valuable as nuclear power. George asserts the first disjunct and Laura the second, so clearly they each trivially accept the entire disjunct. That is simple logic. The problem is that, based on the definition of superiority, intersection goes on to assert the first

⁹An option x is not intersection maximal just when there is some alternative y such that yP_1x holds. Now yP_1x only holds, as per the definition of superiority from section 3.1, when (1) yR_1x is in the intersection R_1 and (2) xR_1y is not in R_1 . (1) obtains provided that for all R_i in $\mathcal{R}$, yR_ix holds. That is, all the value orderings agree that x is either inferior to or equally valuable as some alternative y . (2) obtains provided that there is some ordering R_j from $\mathcal{R}$ that does not assert xR_jy . Since I have been assuming all the orderings in $\mathcal{R}$ are weak orderings, this means that R_j must be connected, and so if xR_jy does not hold then yR_jx must. Therefore, once again by definition of superiority, yP_jx must hold. So (2) holds only when at least one ordering (R_j) judges that option x is inferior to y .

¹⁰The *strong* Pareto principle should not be confused with the *weak* Pareto principle ($\mathcal{P}$) from section 3.2. According to the strong Pareto principle, if (1) for all R_i in $\mathcal{R}$, xR_iy holds and (2) for some R_j in $\mathcal{R}$, xP_jy holds, then xPy must also hold. If the strong Pareto principle holds, then clearly the weak Pareto principle must as well. That is, the antecedent of weak Pareto assumes that for all R_i , xP_iy holds. This clearly satisfies both conditions (1) and (2) of strong Pareto, and so according to strong Pareto xPy holds, which is also the consequent of weak Pareto. Even so, there are ways to obeying weak Pareto without also adhering to strong Pareto, as done by the categorical judgment rule (see footnote 15 in this chapter).

	George	Laura
1 st	coal	coal $\sim$ nuclear
2 nd	nuclear	

FIGURE 5.2: *George and Laura's profile of individual value orderings for choosing between coal power and nuclear power.*

disjunct and reject the second.¹¹ A unanimous judgment on a disjunction suddenly veers into the acceptance of one disjunct, a disjunct that George and Laura do not unanimously accept. Laura's indifference gives George's ordering absolute priority concerning this particular judgment. This is a worrisome bait and switch relying on definitional technicalities. Consequently, intersection maximality makes a weak case for eliminating an option because it may too eagerly remove an option when there is no consensus to do so. George may be eager to eliminate nuclear power, but Laura's ordering suggests she is willing to leave it on the table. That is not much of a consensus.

I also believe that in other social choice contexts blind adherence to strong Paretian impulses raises an additional problem. Imagine a group of one hundred people, where ninety-nine of them agree with Laura, judging that coal power and nuclear power equally valuable. Only George, the one hundredth person, judges that coal power is superior to nuclear power. If the intersection of these judgments represents the collective judgment, then the group ordering ranks coal power over nuclear power. When I presented the money pump argument in section 3.1, I relied on the idea that a judgment of superiority often contains the idea that if an agent judges option x better than option y , it should be willing to 'upgrade' from y to x for some sort of 'transaction fee', the exact amount presumably based on how much better the agent believes x is when compared to y . With this in mind, the problem is that intersection essentially maintains that a collective should be willing to pay for an upgrade when, in fact, only one of its members is willing to do so. Laura and ninety-eight other people judge a move from nuclear power to coal power as 'more of the same' and not as an upgrade. George's judgment to the contrary, however, has intersection judge that the entire group should pay some appropriate price for just such a change, even though an overwhelming majority believes doing so is a waste.

I hasten to add that I am not claiming that George and Laura must then collectively be indifferent between the two options. That would be just as wrong

¹¹Given that George and Laura both can accept the claim that either coal power is superior to nuclear power or coal power is equally valuable as nuclear power, the intersection of their value orderings holds that $(\text{coal})R_1(\text{nuclear})$. Since $(\text{nuclear})R_1(\text{coal})$, on the other hand, does not hold, then according to the definition of superiority, $(\text{coal})P_1(\text{nuclear})$ must hold.

because doing so commits the same problem of accepting a disjunct when there is no consensus to do so. The proper position is, I believe, is that George and Laura should not make any collective judgment linking nuclear power and coal power. This solution is expressed by adopting the following rule for constructing a quasi-ordering from a profile of weak orderings $\mathcal{R}$:

Categorical Judgment: xP_Cy holds if and only if for all value orderings R_i in $\mathcal{R}$, xP_iy holds; and xI_Cy holds if and only if for all value orderings R_i in $\mathcal{R}$, xI_iy holds.¹²

Unlike intersection, categorical judgment only respects unanimous agreements concerning either disjunct and not the entire disjunction. That is, categorical judgment only respects unanimous agreements of the form that ‘option x is superior to option y ’ and agreements of the form ‘option x is equally valuable as option y ’. This distinction between intersection and categorical judgment may be subtle, but it has bite. For instance, the intersection of George and Laura’s value orderings from figure 5.2 asserts that coal power is superior to nuclear power. However, categorical judgment makes no claims linking coal power and nuclear power at all because there is no consensus that the options are equally valuable, nor is there a consensus

¹²While Sen, *On Economic Inequality*, pp. 72–74 suggests acceptance of why I call ‘intersection’, *Inequality Reexamined*, pp. 46–49 seems to suggest what I am calling ‘categorical judgment’ instead. The confusion here may be due to the fact that Sen only applies intersection to profiles $\mathcal{R}$ containing total orderings. A total ordering is weak ordering that also satisfies anti-symmetry: for all options x and y in O , if xR_iy and yR_ix hold then $x = y$. This means that given any two distinct options from O , R_i cannot regard them as equally valuable but must assert that one is superior to the other.

Now it turns out that when $\mathcal{R}$ consists of total orderings, the intersection and categorical judgment rules generate the exact same quasi-ordering. Proving the forward direction, assume that xR_iy holds. Then for all R_i in $\mathcal{R}$, xR_iy holds. Proceeding by cases, assume that $x = y$. By reflexivity, yR_ix holds for all R_i , implying that xI_iy also holds for all R_i . And so xI_Cy holds entailing that xR_Cy must hold as well. Moving to the second case, assume that $x \neq y$. By anti-symmetry, xR_iy and yR_ix cannot both hold for any R_i , and since xR_iy holds for all R_i , yR_ix cannot hold for any R_i . As a result, xP_iy must hold for all R_i . Thus xP_Cy holds, entailing that xR_Cy must hold as well. So much for the forward direction. Showing the backwards direction, assume that xR_Cy holds. This means that either xP_Cy or xI_Cy holds. Proceeding by cases, assume that xP_Cy holds. Then for all R_i in $\mathcal{R}$, xP_iy holds, meaning that xR_iy must hold for all R_i as well. Given that, xR_iy holds. For the second case, assume xI_Cy holds. So for all R_i , xI_iy holds, meaning that xR_iy must hold for all R_i as well. Once again, xR_iy holds. So when $\mathcal{R}$ consists of total orderings, xR_iy holds if and only if xR_Cy holds, meaning that the intersection and categorical judgment rules generate the exact same ordering.

The upshot of all this is that limiting $\mathcal{R}$ to total orderings is one way to avoid intersection’s blind adherence to the strong Pareto principle. However, by requiring anti-symmetry, total orderings bar indifference between different options. This is surely too strong. Applying categorical judgment achieves the same results (that is, when applied to total orderings) but also is capable of defying strong Pareto (even when applied to weak orderings). So rather than revise how the value orderings in $\mathcal{R}$ are understood, it is more direct to abandon intersection in favor of categorical judgment.

that one option is superior to the other. According the quasi-ordering categorical judgment generates in this case, the two options are simply not comparable.¹³

Maximizing according to the quasi-ordering generated by categorical judgment yields a new decision rule:

Categorical Maximization: $M_C(A, \mathcal{R}) = \{x \mid x \in A \text{ and there is no } y \in A \text{ such that } yP_Cx \text{ holds}\}$, where P_C is the categorical judgment generated from $\mathcal{R}$.¹⁴

According to this rule, an option is eliminated when there is an alternative available with unanimous support asserting that alternative's superiority. So for the situation in figure 5.2, both coal and nuclear power are strictly categorically maximal because there is no alternative that George and Laura agree is superior to the other. Thus categorical maximization gives up strong Pareto while still adhering to something like the weak Pareto principle ($\mathcal{P}$) from section 3.2.¹⁵ This is seen with Barry and Michelle from figure 5.1. In that case, categorical maximization eliminates nuclear power because both orderings unanimously agree that coal power is superior to it. This leaves both coal power and hydropower available for choice, just as intersection maximization does. Generally speaking, categorical maximization is never more inclusive than intersection maximization because the former tends to more aggressively eliminate options. That is, any option that is categorically maximal will always be intersection maximal, but not vice versa.¹⁶

As an additional interesting point of contrast with these two maximization rules, intersection may be coupled with optimization to create the following rule:

¹³This provides a different way of understanding the problem with intersection's application in figure 5.2. By asserting that either coal power is superior to nuclear power or coal power is equally valuable as nuclear power, intersection is claiming that the two options are comparable in one of those two ways. Unfortunately, the definitions of superiority and indifference force intersection to specify which one—since they both cannot hold—it actually accepts. Just as constructive logic demands that asserting the truth of a disjunction requires also specifying exactly which disjunct is true, asserting connectivity requires specifying the precise nature of that connection. By denying connectivity, categorical judgment sidesteps that problem.

¹⁴This is identical to what Isaac Levi, *Hard Choices: Decision Making Under Unresolved Conflict* (Cambridge: Cambridge University Press, 1986), p. 94 calls 'V-noninferiority'.

¹⁵Recall that according to $\mathcal{P}$, if for all R_i in $\mathcal{R}$, xP_iy holds, then xPy must also hold. This suggests that if all R_i agree that option x is strictly preferred to option y , then y must not be chosen when x is available. $\mathcal{P}$ is practically enshrined in the definition of categorical maximization.

¹⁶Put more formally, $M_I(A, \mathcal{R})$ is always a subset of $M_C(A, \mathcal{R})$, but not the other way around. For assume that x is not an element of $M_C(A, \mathcal{R})$. Then there is some option y such that yP_Cx holds, meaning that for all R_i in $\mathcal{R}$, yP_ix holds. Consequently, yR_ix holds for all R_i , and xR_iy holds for none of them. This entails that yR_ix holds, but xR_1y does not. Thus yP_1x obtains and so x cannot be an element of $M_I(A, \mathcal{R})$. The converse does not always hold, though. As I have already noted in the text above, when given George and Laura's value orderings from figure 5.2, $M_C(A, \mathcal{R}) = \{\text{coal}\}$ but $M_I(A, \mathcal{R}) = \{\text{coal, nuclear}\}$.

Intersection Optimization: $C_I(A, \mathcal{R}) = \{x \mid x \in A \text{ and for all } y \in A, xR_I y \text{ holds}\}$, where R_I is the intersection generated from $\mathcal{R}$.¹⁷

This selects options that have unanimous support for being optimal according to each and every value ordering in $\mathcal{R}$. For instance, in George and Laura's situation from figure 5.2, both agree that coal power is optimal and so coal power is intersection optimal. There is disagreement over whether nuclear power is also optimal, so intersection optimization eliminates it. Even so, intersection optimization imposes an extremely strict demand on the sort of agreements that must exist between value orderings in order for a decision to be made. So demanding, indeed, that should conflict exist amongst the value orderings in $\mathcal{R}$ concerning the optimal options, there might not be any intersection optimal options. Such is the case with the profile in figure 5.1, for Barry holds that coal power is optimal while Michelle maintains that the optimal choice is hydropower.

It is worth dwelling on this last point for a moment. The demanding nature of Intersection optimization may cause it to flounder in situations of unresolved conflict because it may eliminate all of the available options. In that case, *no* choice is acceptable. I maintain that this makes it a poor decision rule because it offers absolutely no principled guidance in such circumstances when *some* choice must be made.¹⁸ In light of this, optimization can be replaced with the maximization,

¹⁷This is equivalent to what Levi, *Hard Choices*, p. 93 calls 'V-optimality'. Another possible rule would couple optimization with categorical judgment:

Categorical Optimization: $C_C(A, \mathcal{R}) = \{x \mid x \in A \text{ and for all } y \in A, xR_C y \text{ holds}\}$, where R_C the categorical judgment generated from $\mathcal{R}$.

According to this rule, an option x is rejected unless when compared to any alternative y there is either (1) unanimous agreement that x is superior to y or (2) unanimous agreement that x is equally valuable as y . Like that between the two maximization rules, the distinction between intersection and categorical optimization is subtle with bite. Both rules can reject all the available options, as they do when given Barry and Michelle's value orderings from figure 5.1. They can also diverge, as they do when given George and Laura's orderings from figure 5.2, where $C_C(A, \mathcal{R}) = \{\}$ while $C_I(A, \mathcal{R}) = \{\text{coal}\}$.

It turns out that $C_C(A, \mathcal{R})$ is always a subset of $C_I(A, \mathcal{R})$, but not the other way around. For assume that x is an element of $C_C(A, \mathcal{R})$. Then for all y , $xR_C y$ holds. So either $xP_C y$ or $xI_C y$ must hold. Now proceeding by cases, if $xP_C y$ holds for all y , then for all y in A and R_i in $\mathcal{R}$, $xP_i y$ holds. Consequently, for all y and R_i , $xR_i y$ holds, and so $xR_I y$ also holds for all y . Therefore, x is an element of $C_I(A, \mathcal{R})$. Going on to the second case, if $xI_C y$ holds for all y in A , then for all y and R_i in $\mathcal{R}$, $xI_i y$ holds. So for all y and R_i , $xR_i y$ and $yR_i x$ both hold, and so $xR_I y$ also holds for all y . Therefore, x is an element of $C_I(A, \mathcal{R})$. The fact that $C_I(A, \mathcal{R})$ is not always a subset of $C_C(A, \mathcal{R})$ has already been seen by how the rules diverge when looking at George and Laura's situation in figure 5.2. So categorical optimization tends to be even more an aggressive eliminator of options than intersection optimization.

¹⁸Since categorical optimization (see footnote 17 in this chapter) is stricter in its elimination of options than intersection optimization, my concern with the former is more pronounced.

	William	Hillary
1 st	coal	hydro
2 nd	nuclear	nuclear
3 rd	hydro	coal

FIGURE 5.3: *William and Hillary's profile of individual value orderings for choosing from coal power, nuclear power, and hydropower.*

to get intersection maximization. Not only does intersection maximization always identify at least one acceptable alternative, but if there do exist intersection optimal options then these uniquely identify the intersection maximal ones.¹⁹ So this marks a fairly conservative move, but I maintain that such conservatism leaves intersection maximization still too strict because intersection maximization eliminates options that some value orderings may nevertheless consider optimal, as it does in the case of George and Laura. Because of this concern, categorical judgment may replace intersection, with categorical maximality as the result. Compared to the previous two rules, it comes closer to respecting the right sort of agreements between value orderings by only eliminating an option when there is an alternative available that all orderings agree is categorically—literally all things considered—superior to it. This seems to me to provide the correct basis for eliminating options. In what follows, I remain resistant to adopting procedures that cannot respect the weak Pareto principle ($\mathcal{P}$) in this manner.

A concern might be lurking with categorical maximization and weak Pareto's lack of rigor in eliminating options, however. For if there exists severe enough conflict amongst the value orderings, the categorical maximal options may be virtually identical to the available options being chosen from. This is seen, for instance, with William and Hillary's value orderings in figure 5.3. Categorical maximization makes no progress in that situation because all three options are categorically maximal; nothing is eliminated at all. Technically, this is still something, though. When given William and Hillary's orderings, intersection optimization eliminates all options, essentially maintaining that no matter what option is chosen, it is the wrong one. Of course, saying that no matter what option is chosen, it is acceptable, may not be thought as much of an improvement.

Regardless, I believe that when conflict is severe, it is a mistake to assume that a decision procedure alone is enough to miraculously resolve it. Sometimes further

¹⁹This is to say that (1) $C_1(A, \mathcal{R})$ is always a subset of $M_1(A, \mathcal{R})$, and (2) if $C_1(A, \mathcal{R})$ is not empty, then $C_1(A, \mathcal{R}) = M_1(A, \mathcal{R})$. To show (1), assume that x is not an element of $M_1(A, \mathcal{R})$. Then there is some option y such that yP_1x holds, meaning that xR_1y does not hold. Therefore, x cannot be an element of $C_1(A, \mathcal{R})$. Part (2) follows from Lemma 1*^d in Sen, *Collective Choice and Social Welfare*, pp. 11–12.

inquiry and reflection may be necessary to change these value orderings themselves to create more areas of consensus, or methods of conflict resolution may reveal alternative ways to evaluate options or allow for the creation of new options. Such paths may sufficiently change the choice situation, pointing to a narrower set of categorical maximal, or even intersection optimal, options. Doing this is critical for any decision procedure to achieve firmer conclusions; but if extremely severe conflict remains at the moment of choice, however, it should not be surprising that there is little traction for a procedure to seize upon. I return to this concern towards the end of this chapter.

For now, though, I suggest that concerns with the potential indefiniteness of categorical maximization during times of unresolved conflict should not be overstated, for even the default rationality thesis leaves open the possibility that there may be more than one optimal option available for choice. When this occurs, the assumption is that there is some agreed-upon tie-breaking mechanism for ultimately settling on a single option when only one may be chosen. The same certainly applies for choosing amongst the categorically maximal options. One method would be to take these options and select the categorically maximal options amongst them relative to a profile of second-tier value orderings $\mathcal{R}_2$. These would be orderings that were not part of the initial, first-tier orderings in $\mathcal{R}$, now denoted as $\mathcal{R}_1$. To illustrate, suppose that M_{C_1} denotes the set of categorically maximal options determined by profile $\mathcal{R}_1$, that is, $M_{C_1} = M_C(A, \mathcal{R}_1)$. M_{C_2} then denotes those options from M_{C_1} that are also categorically maximal with respect to $\mathcal{R}_2$, that is, $M_{C_2} = M_C(M_{C_1}, \mathcal{R}_2)$. This lexical process of whittling down the options could continue by employing profiles $\mathcal{R}_3$, $\mathcal{R}_4$, and so on, ending when either a single option remaining or all the designated tiers of profiles are exhausted.²⁰

As an example of how this process works, consider the two profiles of value orderings in figure 5.4. According to this setup, building costs and job creation determine the two first-tier value orderings for $\mathcal{R}_1$. In that case, coal power and hydropower are categorically maximal according to these orderings. If only one of these power plants may ultimately be constructed, then $\mathcal{R}_2$ with its value orderings based on environmental impact and maintenance could break this impasse by selecting hydropower as the sole remaining categorically maximal option. It is important to note, though, that $\mathcal{R}_2$ would not have been deployed had $\mathcal{R}_1$ settled on a unique categorically maximal option. This shows how the value orderings in $\mathcal{R}_2$ have very strict, lexical priority over those in $\mathcal{R}_1$. It has been determined that

²⁰This approach is suggested by Levi, *Hard Choices*, pp. 80–82, though there it is employed as an extension to V-admissibility, which I discuss in section 5.2. Nowhere can I find Sen recommending such an approach with categorical maximization, though Sen, ‘Incompleteness and Reasoned Choice’, p. 74 at least hints at approval for it in the context of Levi’s theory.

	$\mathcal{R}_1$		$\mathcal{R}_2$	
	Building Costs	Maintenance Costs	Environmental Impact	Job Creation
1 st	hydro	coal	nuclear	hydro
2 nd	coal	nuclear	hydro	nuclear
3 rd	nuclear	hydro	coal	coal

FIGURE 5.4: Four criteria for ranking coal power, nuclear power, and hydropower are separated into two profiles of value orderings $\mathcal{R}_1$ and $\mathcal{R}_2$, where profile $\mathcal{R}_1$ has lexical priority over $\mathcal{R}_2$.

building costs and job creation are more important than environmental impact and maintenance costs. Even so, there is still no priority or privileged weighing specified for the value orderings within the same tier. So no priority has been specified between building costs and job creation, for instance. So if there is at least some agreement concerning the relative priority of some values, a lexical process of elimination like this can be employed to make a principled decision.

My primary concern with categorical maximality, however, is that it does not make use of inter- or intra-personal information that might exist concerning the value orderings in the same tier.²¹ In that respect, it does not deviate far from the informational constraints following from the default rationality thesis. For instance, according to the value orderings in figure 5.3, William and Hillary do agree that nuclear power ranks second, while disagreeing over which power plant is best and which is worst. As a result, all three options are categorically maximal. However, if interpersonal cardinal information was present concerning these two value orderings, the pair of von Neumann-Morgenstern value functions presented in either profile $\mathcal{V}_*$ or profile $\mathcal{V}^*$ from figure 5.5 could generate the value orderings in figure 5.3.²² However, the differences between these two profiles of value functions appear serious: according to $\mathcal{V}_*$, William and Hillary agree that nuclear power is only slightly better than the option each designates worst, whereas according to $\mathcal{V}^*$, they agree that nuclear power is only slightly worse than the option each designates best. In other words, William and Hillary might agree that nuclear power is the second worst choice (as suggested by $\mathcal{V}_*$) or they

²¹The problem I identify here applies just as well to all the decisions rules I discuss in the chapter when they do not take into consideration neutral probability mixtures.

²²Recall from section 3.1 that it is easy to go from a value function v to a value ordering R by employing the following rule: xRy holds if $v(x) \geq v(y)$ for all options x and y . Using this rule, it is clear how profile $\mathcal{V}_*$ and profile $\mathcal{V}^*$ from figure 5.5 each generate profile $\mathcal{R}$ from figure 5.3. I should emphasize, however, that I am not committing myself here to the claim that every value ordering is the reduction of a value function. Rather, I am only saying that if such cardinal information is available, it is wrong to strip it away and ignore it as categorical maximization does.

might agree that it is the second best choice (as suggested by $\mathcal{V}^*$).²³ Of course, if only the information provided by the orderings in figure 5.3 is available, nuclear power is both. Such a distinction, though, may be important because it signals the difference between a bad compromise, where everyone is worse off by accepting it, and a good compromise, where everyone only sacrifices a little by accepting it. If an option is second worse, and thus a bad compromise, this suggests a compelling reason to eliminate it. An option, on the other hand, that the every value function agrees is second best and a good compromise is certainly something worth further consideration. Categorical maximality, however, ignores this distinction since it simply filters out the relevant interpersonal cardinal information making the distinction meaningful.

It is possible to modify categorical maximality so that it recognizing this distinction. One way of doing this is by requiring that the value orderings in $\mathcal{R}$ not

²³Since these are von Neumann-Morgenstern value functions, the cardinal information they each contain is only unique up to strictly positive affine transformations, John von Neumann and Oskar Morgenstern, *Theory of Games and Economic Behavior*, 3rd ed. (Princeton: Princeton University Press, 1953), p. 627. That is to say, two value functions v and v' contain the exact same cardinal information whenever they are of the form $v'(x) = \alpha + (\beta \times v(x))$, where α and β are real numbers with $\beta > 0$. So, for instance, William's value function, call it v_1 , from profile $\mathcal{V}_*$ in figure 5.5 contains the same cardinal information as v'_1 , where $v'_1(\text{coal}) = 2.50$, $v'_1(\text{nuclear}) = 2.05$, and $v'_1(\text{hydro}) = 2.00$ because v'_1 is a transformation of v_1 , with $\alpha = 2.00$ and $\beta = 0.50$. So there is nothing significant about hydropower being 0.00 according to v_1 , nor does it suggest that, according to v_1 , coal power is ten times better than nuclear power (i.e., because $0.1 \times 10 = 1.0$). All that matters for v_1 is the relative distances—often known as the intervals—between options. It is the interval between coal power and hydropower that must be ten times greater than that between nuclear power and hydropower. Both v_1 and v'_1 agree on these types of claims, and so, insofar as they are von Neumann-Morgenstern value functions, both are functionally identical.

Regardless, this information is sufficient for claiming that William judges that nuclear power is second worst: the interval separating coal power from nuclear power is far greater than that separating nuclear power and hydropower. This suggests that nuclear power is closer in close proximity to hydropower and so not much better than it. According to William's function in $\mathcal{V}^*$, on the other hand, nuclear power is second best because that first interval is now far less than the second one. In that case, nuclear power appears in close proximity to coal power and so not much worse than it.

Less important, but still worth noting, is that I am assuming that this cardinal information concerning intervals also extends to interpersonal comparisons. This means that the intervals going across different value functions in are also relevant. For instance, this interpersonal information states that according to $\mathcal{V}_*$ the interval between William's assessment of coal power and Hillary's assessment of nuclear power is nineteen times greater than that separating William's assessment of nuclear power from Hillary's assessment of coal power. So just like the information concerning the individual value functions, the entire table in figure 5.5 representing $\mathcal{V}_*$ is also unique to up to strictly positive affine transformations. Transforming v_1 to v'_1 as I did above does preserve William's *cardinal* information, but without also applying the same transformation (with $\alpha = 2.00$ and $\beta = 0.50$) to Hillary's value function, loses their shared *interpersonal* information because a completely different set of interpersonal cardinal intervals would result.

	$\mathcal{V}_*$		$\mathcal{V}^*$	
	William	Hillary	William	Hillary
coal	1.0	0.0	1.0	0.0
nuclear	0.1	0.1	0.9	0.9
hydro	0.0	1.0	0.0	1.0

FIGURE 5.5: Two profiles of von Neumann-Morgenstern value functions representing alternative ways in which William and Hillary from figure 5.3 could evaluate the options numerically.

only rank the available options $\mathcal{A}$ but also rank all the neutral probability mixtures of those options.²⁴ If $\text{Mix}(\mathcal{A})$ denotes the set consisting of the available options $\mathcal{A}$ plus all neutral probability mixtures involving them, categorical maximality can be extended as follows:

Categorical Maximization Over Mixtures: $M_M(\mathcal{A}, \mathcal{R}) = \{x \mid x \in \mathcal{A} \text{ and there is no } y \in \text{Mix}(\mathcal{A}) \text{ such that } y P_C x \text{ holds}\}$, where P_C is the categorical judgment generated from $\mathcal{R}$.²⁵

So for instance, given the situation in figure 5.3, a neutral probability mixture of coal power and hydropower might involve tossing a fair coin to choose between them, so long as the respective judgments of William and Hillary are not influenced by an attraction towards or an aversion against settling things in this way. This mixed option may be denoted as $0.5(c) \oplus 0.5(h)$. If this option is considered along with coal power, nuclear power, and hydropower, then the profile of value functions $\mathcal{V}_*$ from figure 5.5 generates the profile of value orderings $\mathcal{R}_*$ presented in figure 5.6.²⁶ Those orderings do reveal that nuclear power is indeed second worst, because William and Hillary now agree that they are better off by using the

²⁴A probability mixture selects an option from a pre-specified (finite) set of options, where each option has some specified chance of being selected. Put more formally, a probability mixture may be understood as a lottery involving options $x_1, x_2, \dots, x_m$ and nonnegative numbers $p_1, p_2, \dots, p_m$ summing to 1, where option x_j has the objective statistical probability p_j of being selected by that lottery. Such a mixture may be denoted by $p_1 x_1 \oplus p_2 x_2 \oplus \dots \oplus p_m x_m$. So instead of directly choosing an option, an agent might consider instead indirectly choosing an option through the implementation of a probability mixture. A probability mixture is neutral with respect to a value ordering or value function provided that its associated value is not influenced by an aversion or desire for randomly selecting options. In particular, a probability mixture is neutral provided that the value of choosing any one of the options it involves is the same regardless of whether the option was directly chosen or chosen as the result of implementing that mixture. For more on neutral probability mixtures, see Levi, *Hard Choices*, pp. 74–77.

²⁵Extending intersection maximization in this way would also allow it to distinguish between second best and second worst, though its adherence to strong Pareto would still remain.

²⁶The fact that the value functions in $\mathcal{V}_*$ and $\mathcal{V}^*$ from figure 5.5 are von Neumann-Morgenstern functions is now relevant here. According to such functions, the value of a neutral probability

coin to randomly choose between the other two power plants than going with nuclear power. Consequently, nuclear power is not categorically maximal when mixtures are considered. On the other hand, profile $\mathcal{V}^*$ generates the profile $\mathcal{R}^*$, where William and Hillary now agree that nuclear power is superior to the coin toss, illustrating how it is both second best and categorically maximal even when mixtures are included.²⁷

This illustrates the extreme sensitivity of categorical maximization concerning whether it considers neutral probability mixtures. Without doing this, however, it cannot meaningfully distinguish between second best and second worst op-

mixture is equal to its expected value:

$$v(p_1x_1 \oplus p_2x_2 \oplus \dots \oplus p_mx_m) = \sum_{k=1}^m (p_k \times v(x_k)).$$

So when the coin flip between coal power and hydropower is a neutral probability mixture, all the value functions in $\mathcal{V}_*$ and $\mathcal{V}^*$ agree in giving this mixture a value of 0.50. Therefore this coin toss is superior to nuclear power according to all the value functions in $\mathcal{V}_*$, while it inferior to nuclear power according to the value functions in $\mathcal{V}^*$.

Now, if either William or Hillary had an attraction or aversion to using a fair coin flip to make the decision, then this mixture would not be neutral and this property of von Neumann-Morgenstern functions would not apply, allowing the value of the coin flip to diverge from its expected value. This explains why the repeated objection to equating value with expected value in Amartya K. Sen, 'Rationality and Uncertainty', *Theory and Decision* 18, 2 (March 1985): 109–127, pp. 119–120, and 'Maximization and the Act of Choice', pp. 758–759, does not apply here. According to that argument, suppose Dr. Chang has two critically ill patients, Hao and Lin, but only enough antidote to administer to one of them. If faced with the direct choice of who gets the antidote, Dr. Chang favors giving it to Hao because Dr. Chang believes that the antidote is 5 percent more likely to cure Hao than to cure Lin. Even so, Dr. Chang might prefer even more to simply let a fair coin toss decide who gets the cure, creating the mixed option $0.5(\text{give to Hao}) \oplus 0.5(\text{give to Lin})$, which Dr. Chang believes has the most value. This is a violation of the von Neumann-Morgenstern independence postulate concerning mixtures. According to axiom (3:B:a) in Neumann and Morgenstern, *Theory of Games and Economic Behavior*, p. 26, if option x is superior to option y , then x should also be superior to $0.5x \oplus 0.5y$. Dr. Chang's judgments do not conform to this. However, this case presents a violation precisely because Dr. Chang is not neutral with respect to the coin flip: he believes the value of directly choosing a treatment is worse than it being chosen as a result of flipping the coin. In this particular case, at least, Dr. Chang has an attraction towards gambling.

²⁷Strictly speaking, to claim that nuclear power is categorically maximal in this case, it has to be verified that according to the value functions in $\mathcal{V}^*$ there is no neutral probability mixture $\mathcal{M} = p_1(\text{coal}) \oplus p_2(\text{nuclear}) \oplus (1 - p_1 - p_2)(\text{hydro})$, for some nonnegative p_1 and p_2 whose sum does not exceed 1, that William and Hillary agree is superior to nuclear power. In fact, there is no such mixture $\mathcal{M}$. First, the quantity $a = p_1 + (0.1 \times p_2)$ must be less than 0.1 for Hillary to find $\mathcal{M}$ superior to nuclear power. Second, for William to agree that $\mathcal{M}$ is better than nuclear power, a must be greater than $b = 0.9 - (0.8 \times p_2)$. But satisfying both requirements is impossible because for b to be less than 0.1, p_2 would have to be greater than 1, which is not possible for $\mathcal{M}$. Consequently, there is no neutral probability mixture $\mathcal{M}$ that William and Hillary agree is superior to nuclear power. So nuclear power remains categorically maximal even when all neutral mixtures are considered.

	$\mathcal{R}_*$		$\mathcal{R}^*$	
	William	Hillary	William	Hillary
1 st	coal	hydro	coal	hydro
2 nd	$0.5(c) \oplus 0.5(h)$	$0.5(c) \oplus 0.5(h)$	nuclear	nuclear
3 rd	nuclear	nuclear	$0.5(c) \oplus 0.5(h)$	$0.5(c) \oplus 0.5(h)$
4 th	hydro	coal	hydro	coal

FIGURE 5.6: *Two profiles of value orderings representing alternative extensions of the profile from figure 5.3 taking into consideration the neutral probability mixture where coal power is selected with a 50 percent chance or hydropower is selected with a 50 percent chance.*

tions, which I maintain is an important consideration for eliminating options in addition to weak Pareto. Furthermore, if the value orderings in $\mathcal{R}$ rank neutral probability mixtures in accordance with the von Neumann-Morgenstern postulates, then categorical maximality over mixtures becomes virtually identical in its prescriptions to those made by what Isaac Levi calls ‘V-admissibility’.²⁸ As a result, I now turn towards this account.

5.2 Admissibility

Isaac Levi presents a theory of rational choice under unresolved conflict that explicitly takes into consideration the various forms of information that may or may not be available to decision makers.²⁹ According to this framework, when there is conflict between rival assessments of options, and choice must reflect a commitment to each of these assessments, making a decision requires suspending all judgments concerning which assessments are the right ones. Instead, the decision maker is supposed to consider all of them as equally permissible. In a novel

²⁸The three postulates dictate that the rankings satisfy the requirements of (1) a weak ordering; (2) the independence axiom: for all options x, y , and z with $0 < p \leq 1$, xRy holds if and only if $(px \oplus (1-p)z)R(px \oplus (1-p)z)$ holds; and (3) the Archimedean axiom: if xPy and yPz , then there are some p_1 and p_2 where $0 < p_1, p_2 < 1$, such that $(p_1x \oplus (1-p_1)z)Py$ and $yP(p_2x \oplus (1-p_2)z)$ hold. For their specifics, see Neumann and Morgenstern, *Theory of Games and Economic Behavior*, pp. 26–29. For discussion on the relationship between categorical maximality and Levi’s account of V-admissibility, see Sen, ‘Incompleteness and Reasoned Choice’, pp. 49–50; Levi, *Hard Choices*, pp. 91–95, and ‘Amartya Sen’, *Synthese* 140, 1–2 (May 2004): 61–67, pp. 62–63, and footnote 35 in this chapter.

²⁹‘Conflict and Social Agency’, *Journal of Philosophy* 79, 5 (May 1982): 231–247; *Hard Choices*; ‘Pareto Unanimity and Consensus’, *Journal of Philosophy* 87, 9 (September 1990): 481–492; ‘Conflict and Inquiry’, *Ethics* 102, 4 (July 1992): 814–834; ‘Value Commitments, Value Conflict, and the Separability of Belief and Value’, *Philosophy of Science* 66, 4 (December 1999): 509–533; ‘Inclusive Rationality’, *Journal of Philosophy* 101, 5 (May 2004): 255–276; and ‘Amartya Sen’.

move, Levi suggests that for any pair of permissible—though possibly conflicting—assessments, all the possible resolutions and compromises between them should also be considered as permissible assessments of the options. Finally, an option is admissible for choice provided that at least one permissible assessment regards it as optimal, that is, as superior to all the other available options.³⁰ Conversely, should consensus amongst all permissible evaluations be that a particular option is not optimal, that option is inadmissible. This reflects the type of agreement that Levi believes a decision should respect: unanimous agreement of non-optimality implies that an option should not be chosen. So while conflicting evaluations may disagree over what is optimal, they may agree on what is not. As I suggested at the beginning of this chapter, I believe that this general approach respects the right sort of agreements that should be considered for eliminating options, though I have concerns with how Levi's particular implementation runs afoul of weak Pareto and other considerations. I return to these at the end of this section.

Regardless, to illustrate Levi's approach, consider William and Hillary's situation from figure 5.3. If a decision here is supposed to reflect a commitment to each of these rival orderings, then William and Hillary are supposed to suspend all judgments concerning who is right and who is wrong. Both orderings are therefore permissible, thus making both coal power and hydropower admissible options. The question is whether nuclear power is also admissible. If these two orderings are the only permissible evaluations, then nuclear power is inadmissible. However, Levi's theory advises William and Hillary to consider all the potential compromises between that exist between their orderings. Nuclear power may then be admissible, provided that it is regarded as optimal by at least one of these compromises.

Rather than abandon William and Hillary to work out the compromises on their own, Levi maintains that the potential compromises are determined by information concerning how William and Hillary arrive at their respective evaluations. That is, Levi has William and Hillary explore the reasons behind their respective orderings and discover what sorts of agreements and conflicts concerning these reasons they possess, building potential resolutions and compromises based on these. If only this ordering information is available, however, Levi suggests that the potential compromises are simply all the weak orderings consistent with the quasi-ordering produced by the categorical judgment rule.³¹

³⁰The assumption here is that a single mode of evaluation should satisfy the requirements of a weak ordering so that it may identify an optimal option. If an evaluation is not like this, I suspect that Levi would suggest breaking down that evaluation into further, distinct evaluations, each of which is itself a weak ordering. This certainly seems in harmony with the Levi's approach of exploring what underlies or goes into an evaluation.

³¹At least this is my understanding of Levi, 'Amartya Sen', p. 62. The framework in *Hard Choices*,

The idea here is that William and Hillary should consider all the possible ways of weakly ordering the options that is consistent with their shared, categorical judgments. An option is admissible provided it is optimal according to at least one of these weak orders. It turns out, though, that this notion of admissibility identifies exactly the same options as categorical maximization (without mixtures).³² So coal power, nuclear power, and hydropower all satisfy this definition of admissibility.

While Levi's approach may be similar to Sen's at this stage, it quickly begins to diverge as the reasons giving shape to the value orderings are explored. For instance, if William and Hillary's orderings are based on the interpersonal cardinal information given by the von Neumann-Morgenstern value functions in $\mathcal{V}_*$ from figure 5.5, Levi argues that the only permissible compromises between William and Hillary's respective evaluations are all the possible weighted averages of their two value functions. The idea here is that their decision should consider all the possible ways of weighing the importance of each of their individual judgments. Some such ways accord William's value function more and more importance while others do the same for Hillary's. Levi believes that beyond the value functions resulting from these weighted averages, there is nothing else to consider. All such value functions are presented in figure 5.7. As seen there, at no point does nuclear power rise above both the other two options, implying that it is never an optimal option and so it is inadmissible.

The situation is different, however, if William and Hillary's rankings come from the value functions in $\mathcal{V}^*$ from figure 5.5. All the possible value functions resulting from weighted averaging appear in figure 5.8. Now nuclear power does rise above the other two options in the middle area, revealing that it is optimal according to some permissible evaluations, such as when William and Hillary's respective value functions receive equal weight (seen in the area on the graph halfway between the left and right vertical bars). Thus nuclear power is admissible in this case. As a result, Levi's account of admissibility is directly responsive to the distinction between second worst and second best options. A second best option really is a good compromise because it will rise to the top of some permissible resolution to the conflict. A second worst option fails to do this; none of the possible compromises suggest choosing it.

pp. 84–95, concerning categorical preference is somewhat idiosyncratic because it treats relations P and I as primitives rather than as the usual derivations from R . According to the framework I am employing, a weak ordering R_W is consistent with a quasi-ordering R_Q provided that if $xR_Q y$ holds then $xR_W y$ also holds. The assumption here is that there is always at least one weak ordering that is consistent with a quasi-ordering produced by the categorical judgment rule, which is, in fact, a consequence of Theorem 3 in Suzumura, 'Remarks on the Theory of Collective Choice', p. 387.

³²This is a consequence of Proposition 3.2 in Asis Banerjee and Prasanta K. Pattanaik, 'A Note on a Property of Maximal Sets and Choice in the Absence of Universal Comparability', *Economics*

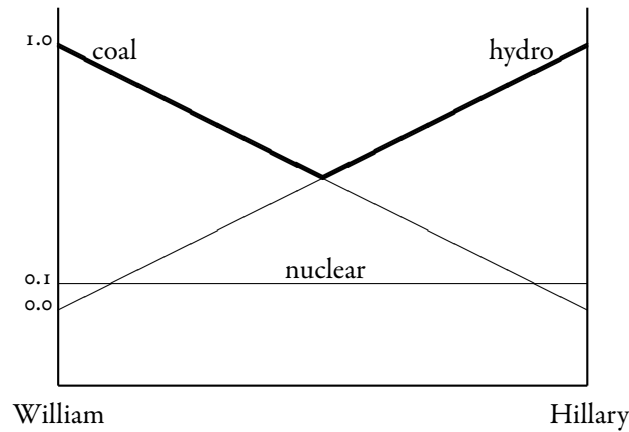


FIGURE 5.7: The graph of how each weighted average of the two value functions in $\mathcal{V}_*$ from figure 5.5 evaluates coal power, nuclear power, and hydropower. The left vertical bar represents giving all weight to William's value function, while the right vertical bar represents doing so for Hillary's. The thick line segments show where an option is optimal according to a weighted average. There is no point at which nuclear power is optimal.

This form of admissibility, operating on a set of von Neumann-Morgenstern value functions $\mathcal{V} = \langle v_1, v_2, \dots, v_n \rangle$, encapsulates what Levi calls 'V-admissibility'. It may be formalized as follows:

V-Admissibility: $A_V(A, \mathcal{V}) = \{x \mid x \in A \text{ and there exists a value function } v \text{ constructed from } \mathcal{V} \text{ by the weighted average principle such that for all } y \in A, v(x) \geq v(y)\}$.³³

The Weighted Average Principle: A value function v is constructed from $\mathcal{V}$ by the weighted average principle by taking any nonnegative numbers $w_1, w_2, \dots, w_n$ summing to 1 and setting

$$v(x) = \sum_{i=1}^n (w_i \times v_i(x)), \text{ for each option } x.^{34}$$

Letters 51, 2 (May 1996): 191–195, p. 192.

³³Levi, *Hard Choices*, pp. 68–79.

³⁴Being able to perform this summation and multiplication assumes that the value functions in $\mathcal{V}$ are cardinally inter- or inter- or intra-personally comparable, for it would not make sense to add incomparable quantities together. See *ibid.*, pp. 77–79, and Levi, 'Amartya Sen', for further

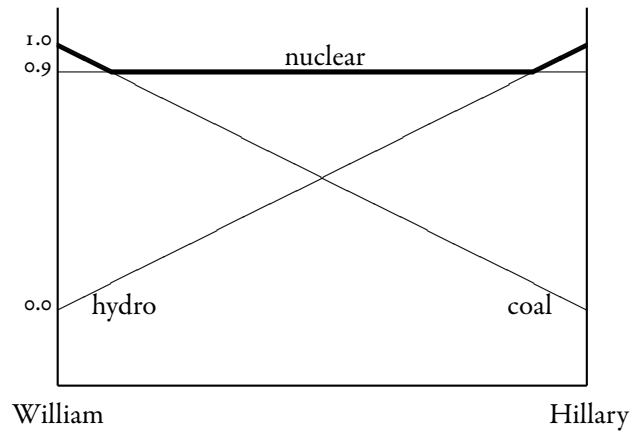


FIGURE 5.8: The graph of how each weighted average of the two value functions in $\mathcal{V}^*$ from figure 5.5 evaluates coal power, nuclear power, and hydropower. The left vertical bar represents giving all weight to William's value function, while the right vertical bar represents doing so for Hillary's. The thick line segments show where an option is optimal according to a weighted average. In this case, there are points at which nuclear power is optimal.

In sum, V-admissibility looks at every possible weighted average of the value functions in $\mathcal{V}$ and selects just those options that rank at the top of at least one of those evaluations. In this case, it turns out that this conception of admissibility identifies the same options as categorical maximization over mixtures does, provided that the corresponding value orderings for maximization rank neutral probability mixtures in accordance with the von Neumann-Morgenstern postulates.³⁵ Furthermore,

explication of this principle. If there is no disagreement between how the value functions in $\mathcal{V}$ rank the options in $\mathcal{A}$, then the weighted average principle has little operational content. That is, if two value functions have the same ranking of options, then any weighted average of them also has that same ranking of options. However, if there are conflicting assessments in $\mathcal{V}$, then it is important to recognize that each value function v_i in $\mathcal{V}$ is itself a potential resolution. Using the weighted average principle, this simply means giving all the weight to v_i (by setting $w_i = 1$) and none to the other functions in $\mathcal{V}$ (by setting $w_j = 0$ for all $j \neq i$).

³⁵In addition to the stated assumptions, I am also assuming that the value functions in $\mathcal{V}$ order the options the same way as the respective value orderings in $\mathcal{R}$, i.e., $v_i(x) \geq v_i(y)$ if and only if xR_iy for all x and y . Now a consequence of Lemma 3 from David G. Pearce, 'Rationalizable Strategic Behavior and the Problem of Perfection', *Econometrica* 52, 4 (July 1984): 1029–1050, p. 1048, rephrased more clearly as Theorem 1 in Teddy Seidenfeld, Mark J. Schervish and Joseph B. Kadane, 'Coherent Choice Functions Under Uncertainty', *Synthese* 172, 1 (January 2010): 157–176, p. 160, is that if an option x is not in $\mathcal{A}_V(\mathcal{A}, \mathcal{V})$, then there must be some neutral probability mixture y such that $v_i(y) > v_i(x)$ for all v_i . Consequently, $yP_i x$ holds for all R_i , and so $yP_C x$ holds. Thus x is not in $M_M(\mathcal{A}, \mathcal{R})$.

V-admissibility diverges from the standard accounts of revealed preference because its prescriptions do not reduce to pairwise comparisons. That is, how V-admissible ranks every pair of options may not reveal how it chooses amongst larger sets of options.³⁶

Even so, the gap between these two approaches increases because admissibility is not limited just to von Neumann-Morgenstern value functions. For instance, William and Hillary's value orderings in figure 5.3 could instead be based upon Bayesian value functions incorporating the uncertainties involved in building a power plant.³⁷ Recall from section 3.3 that a Bayesian value function v can be

Before showing the opposite direction, first recognize that if aRb_1 and aRb_2 both hold, then for all p such that $0 < p < 1$, $aR(pb_1 \oplus (1-p)b_2)$ holds. For according to the independence postulate on neutral probability mixtures (see footnote 28 in this chapter), aRb_2 entails that for all p , $0 < p < 1$, $aR(pa \oplus (1-p)b_2)$ holds, while aRb_1 entails that for all p , $0 < p < 1$, $(pa \oplus (1-p)b_2)R(pb_1 \oplus (1-p)b_2)$ holds. And so by transitivity of R , $aR(pa \oplus (1-p)b_2)$ holds. The contrapositive of this mixture property can then be generalized to claim that if there exist options $b_1, b_2, \dots, b_m$ and nonnegative numbers $p_1, p_2, \dots, p_m$ summing to 1 where $(p_1b_1 \oplus p_2b_2 \oplus \dots \oplus p_mb_m)Pa$ holds, then b_kPa holds for at least one of the b_k 's. More simply put, if a mixture is superior to an option a , then at least one of the options involved in that mixture should be better than a .

Now assume that x is not in $M_M(A, \mathcal{R})$. Then there is some y such that yP_Cx holds, and so for all R_i in $\mathcal{R}$, $yP_i x$ holds. Now if y is not a neutral probability mixture, then y is in A and $v_i(y) > v_i(x)$ for all v_i in $\mathcal{V}$, and so x cannot be in $A_V(A, \mathcal{V})$. On the other hand, if y is a neutral probability mixture, then $y = (p_1y_1 \oplus p_2y_2 \oplus \dots \oplus p_my_m)$ for some options $y_1, y_2, \dots, y_m$ in A and nonnegative numbers $p_1, p_2, \dots, p_m$ summing to 1. This means for each R_i , there is some y_k , such that $y_kP_i x$ holds. So $v_i(y_k) > v_i(x)$ for all v_i , implying that x cannot be in $A_V(A, \mathcal{V})$.

³⁶One such condition of for reducing judgment to pairwise comparisons is Sen's property γ , which states that for any collection of finite subsets A_i of options in O , if x is chosen when given any one of these A_i , then x must also be chosen from the union of all the A_i 's, Amartya K. Sen, 'Choice Functions and Revealed Preference', *Review of Economic Studies* 38, 3 (July 1971): 307–317, p. 314. This means, for instance, that if x is chosen from all the possible pairs of options involving it, then it should also be chosen when all the options are available. V-admissibility violates this, however, as the value functions in profile $\mathcal{V}_*$ from figure 5.5 illustrate. For when given just nuclear power and coal power, nuclear power is V-admissible (as is coal power), and when given just nuclear power and hydropower, nuclear power is again V-admissible (as is hydropower). But when given all three options, nuclear power is no longer V-admissible. So the fact that nuclear power is chosen in all the pairwise comparisons does not reveal whether it is chosen from the three-way comparison, which depends instead on whether it is second worst or second best relative to the other two options.

Another condition of revealed preference violated by V-admissibility is Sen's property β , which states that if A is a subset of options in B and options x and y are chosen from A , then x is chosen from B as long as y is also chosen from B , *ibid.*, p. 313. Put informally, winners (x and y) cannot be split by increasing the set of available options (going from A to B). To see how V-admissibility violates this, suppose that $\mathcal{V}$ consists of v_1 and v_2 , where $v_1(\text{coal}) > v_1(\text{nuclear})$, $v_1(\text{nuclear}) = v_1(\text{hydro})$, $v_2(\text{coal}) = v_2(\text{nuclear})$, and $v_2(\text{nuclear}) > v_2(\text{hydro})$. In that case, when given nuclear power and hydropower, both options are the V-admissible 'winners', but when given all three options, nuclear power remains such a 'winner' while hydropower is not. The winners have been split by making coal power available.

³⁷What I am calling 'Bayesian' value functions concern the evaluation of Anscombe-Aumann

broken into judgments concerning the probability and evaluations of the possible consequences, where the value of an option is identical to its expected value given all its potential consequences.³⁸ So suppose that the conflict between William and Hillary arises from a disagreement concerning whether the local mines will continue to produce coal. William believes that it is likely that there is plenty of coal for the foreseeable future, whereas Hillary believes that it is more likely that the mines are tapped out. Both agree, however, that if the coal runs out, then coal power is the worst option, while hydropower is the best because it would provide the most job opportunities for unemployed miners. They also agree that should the mines remain productive, coal power is the best option, and hydropower is the worst because the potential labor pool would be insufficient to meet the massive demands of constructing and maintaining a hydropower dam. Finally, they agree that either event has no effect on the relative value of nuclear power, which they both agree is fairly small. Suppose that the decision matrix in figure 5.9 provides the values of each these various possibilities, where ω_1 denotes the state of affairs

'horse lotteries', F.J. Anscombe and Robert J. Aumann, 'A Definition of Subjective Probability', *Annals of Mathematical Statistics* 34, 1 (March 1963): 199–205. These are in contrast to von Neumann-Morgenstern value functions, which concern the evaluation of neutral probability mixtures in accordance with the von Neumann-Morgenstern postulates (see footnotes 24, 26, and 28 in this chapter). According to a neutral probability mixture, there is no question concerning the probability of the outcomes. Flipping a fair coin to choose between options x and y , for instance, is commonly thought to involve an 'objective' probability of a fifty percent chance of selecting either option. So a von Neumann-Morgenstern value function involves assessments of lotteries where the probabilities are pre-given in such a fashion. Horse lotteries are different in two important respects. Recalling the decision matrix in figure 3.3 from section 3.3, (1) the probabilities concerning the states of affairs $\Omega = \langle \omega_1, \omega_2, \dots, \omega_p \rangle$ are not specified in advance for the decision maker, analogous to the chances that a given horse will win a particular race (hence the name of these type of lotteries), and (2) each consequence $c_{k,j}$ is itself a neutral probability mixture with pre-given probabilities, though such a mixture may be the 'sure-thing' that a given outcome occurs with one-hundred percent assurance (most of my examples in this section involve consequences of this sure-thing variety). This means that a Bayesian value function takes into account both 'objective' (specified in advance) and 'subjective' (not so specified) uncertainty. The evaluation of horse lotteries by a Bayesian value function obeys conditions similar to the von Neumann-Morgenstern postulates, though with the important addition that judgments of the neutral probability mixtures are state-independent. See *ibid.*, pp. 200–201.

³⁸That is, according to a Bayesian value function v , for any option x ,

$$v(x) = \sum_{j=1}^p (\Pr(\omega_j) \times \bar{v}(c_{x,j})).$$

$\Pr$ is a probability distribution over the states of affairs in Ω , $\Pr(\omega_j)$ is the probability that state ω_j holds, $\bar{v}$ is the von Neumann-Morgenstern value function over all possible consequences of implementing any option, and $\bar{v}(c_{x,j})$ is the numerical valuation of the consequences for implementing option x when state ω_j holds.

	ω_1	ω_2
coal	1.5	-0.5
nuclear	0.1	0.1
hydro	-0.5	1.5

FIGURE 5.9: A decision matrix representing the uncertainties involved in building a power plant. In choosing from coal power, nuclear power, and hydropower, there are two possible states of affairs with all the potential consequences having the numeric values presented in the table above.

where the mines continue to produce coal and ω_2 is the state of affairs where they do not.

In this case, the conflict between William and Hillary does not concern the value of the various possible consequences, but their respective beliefs concerning the probabilities of ω_1 and ω_2 . Suppose that William assigns probability 0.75 to ω_1 and 0.25 to ω_2 while Hillary gives 0.25 to ω_1 and 0.75 to ω_2 . Given these beliefs and the values in the decision matrix, William and Hillary's respective Bayesian value functions generate the exact same interpersonal cardinal assessments as those given in $\mathcal{V}_*$ from figure 5.5.³⁹ In a similar manner as done there, Levi suggests that the permissible resolutions and compromises to this conflict consist entirely of all the Bayesian value functions formed from taking the weighted averages of the two conflicting probability functions. There is no need to do any averaging of their judgments concerning the possible consequences because William and Hillary agree on the evaluations given in the decision matrix. In particular, this means that the permissible compromises are those Bayesian value functions calculating the expected value of each power plant for probability ω_1 within the interval 0.75 to 0.25 and probability $\omega_2 = 1 - \omega_1$. All these value functions are presented in figure 5.10, demonstrating that coal power and hydropower are admissible, while nuclear power is second worst and not admissible. Had the values of nuclear power been 0.9 instead of 0.1 under both ω_1 and ω_2 , however, then it would have been second best admissible, for reasons analogous to those for its V-admissibility in the case of $\mathcal{V}_*$ from figure 5.5.⁴⁰

³⁹I should emphasize, however, that the pair of value functions in $\mathcal{V}_*$ are *von Neumann-Morgenstern* value functions, and not *Bayesian* value functions (for the distinction see footnote 37 in this chapter), because they do not explicitly concern the uncertainties involved in constructing a power plant. Even so, in this particular case, I have specified William and Hillary's respective probability assessments so that it turns out that the values in the decision matrix from figure 5.9 generate a pair of Bayesian value functions assigning values to the three power plants that are identical to those given by the pair of von Neumann-Morgenstern value functions.

⁴⁰Peter Walley, *Statistical Reasoning with Imprecise Probabilities* (London: Chapman and Hall, 1991), closes this gap between categorical maximality and admissibility by formulating a version of

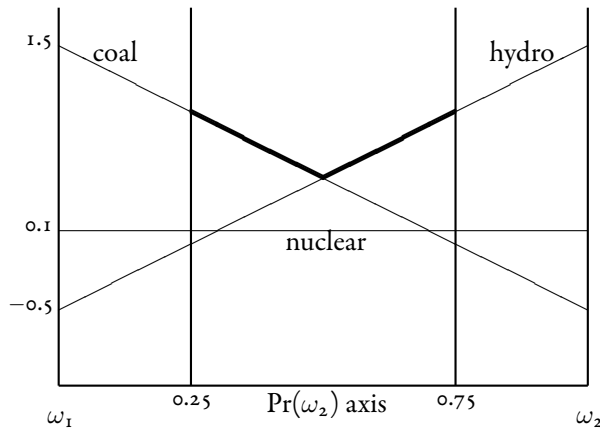


FIGURE 5.10: The graph of the expected value of adopting coal power, nuclear power, and hydropower given the values from the decision matrix in figure 5.9. The left vertical bar represents assigning probability 1 to ω_1 (and 0 to ω_2), while the right vertical bar represents assigning probability 1 to ω_2 (and 0 to ω_1). The thick line segments show where an option is optimal according to a weighted average. There is no point at which nuclear power is optimal.

More generally, Levi holds that when two Bayesian value functions conflict, a potential compromise consists of a Bayesian value function constructed from coupling any weighted average of the rival probability distributions with any weighted average of the rival evaluations of the consequences. Recalling the terminology from section 3.3, this can be put more formally as a decision rule operating on profiles of probability/value pairs $\mathcal{PV} = \langle (\Pr_1, \bar{v}_1), (\Pr_2, \bar{v}_2), \dots, (\Pr_n, \bar{v}_n) \rangle$ for identifying the ‘E-admissible’ options from $\mathcal{A}$:

E-Admissibility: $\mathcal{A}_E(\mathcal{A}, \mathcal{PV}) = \{x \mid x \in \mathcal{A} \text{ and there exists a Bayesian value function } v \text{ constructed from } \mathcal{PV} \text{ by the extended weighted average principle such that for all } y \in \mathcal{A}, v(x) \geq v(y)\}$.⁴¹

The Extended Weighted Average Principle: A Bayesian value function v con-

the former that operates in situations of uncertainty. When such an account of maximality takes into account neutral probability mixtures obeying the von Neumann-Morgenstern postulates, it also begins to converge on Levi’s E-Admissibility along lines similar to those presented in footnote 35 of this chapter. For more, see Mark J. Schervish, Teddy Seidenfeld, Joseph B. Kadane and Isaac Levi, ‘Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons’, in *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, ed. by Jean-Marc Bernard, Teddy Seidenfeld and Marco Zaffalon (Carleton Scientific, 2003): 496–510; and Seidenfeld, Schervish and Kadane, ‘Coherent Choice Functions Under Uncertainty’.

⁴¹Levi, *Hard Choices*, pp. 109–122.

structed from $\mathcal{PV}$ by the extended weighted average principle has the following form:

$$v(x) = \sum_{j=1}^p (\text{Pr}(\omega_j) \times \bar{v}(c_{x,j})), \text{ for each option } x.$$

Here Pr is the result of some weighted average of the Pr_i 's from $\mathcal{PV}$ and $\bar{v}$ is the result of some weighted average of the $\bar{v}_i$'s from $\mathcal{PV}$.⁴²

The motivation here is similar to that for V-admissibility: a potential resolution to a conflict involves weighted averaging. Applying this idea to E-admissibility, Levi advocates separating Bayesian value functions into their probability and evaluational components and seeking out potential compromises, via weighted averaging, for each of these two components individually. Once these two sets of compromises—one concerning the probability distributions and the other concerning the von Neumann-Morgenstern value functions evaluating the possible consequences—are identified, then a single set of Bayesian value functions is constructed from all the possible ways of combining one element from each of the other two sets. These value functions then isolate the E-admissible options.

To illustrate, suppose that William and Hillary disagree over the probabilities of ω_1 and ω_2 as they did in the previous example. However, they now disagree over the values of the potential consequences. William still sees evaluates the consequences as given in figure 5.9. Hillary, on the other hand, now believes that the continued productivity of the coal mines would actually make nuclear power more appealing than coal power because exporting the coal is more profitable than burning it, and these profits would more than offset the costs of building a more environmentally friendly nuclear reactor. However, if the mines run out, then Hillary is less concerning about the environment, and so maintains that coal power is slightly better than nuclear because its associated costs are less and coal can be

⁴²Determining a weighted average of the Pr_i 's or the $\bar{v}_i$'s is similar to that done by the weighted average principle for V-admissibility. So a probability distribution Pr is constructed from $\mathcal{PV}$ by taking any nonnegative numbers $w_1, w_2, \dots, w_n$ summing to 1 and setting

$$\text{Pr}(\omega) = \sum_{i=1}^n (w_i \times \text{Pr}_i(\omega)), \text{ for each state of affairs } \omega.$$

Likewise, a value function $\bar{v}$ is constructed from $\mathcal{PV}$ by taking any nonnegative numbers $\bar{w}_1, \bar{w}_2, \dots, \bar{w}_n$ summing to 1 and setting

$$\bar{v}(c) = \sum_{i=1}^n (\bar{w}_i \times \bar{v}_i(c)), \text{ for each consequence } c.$$

	William		Hillary	
	ω_1	ω_2	ω_1	ω_2
coal	1.5	-0.5	0.9	-0.3
nuclear	0.1	0.1	1.6	-0.4
hydro	-0.5	1.5	0.4	1.2

FIGURE 5.11: *Two decision matrices representing the rival ways in which William and Hillary evaluate all the potential consequences of building a power plant.*

imported from the surrounding region. William and Hillary's respective decision matrices for this situation are depicted in figure 5.11. Since William assigns probability 0.75 to ω_1 while Hillary gives it 0.25, these beliefs and evaluations also generate the exact same interpersonal cardinal assessments as those given in profile $\mathcal{V}_*$ from figure 5.5.⁴³ In this case, however, it turns out that in addition to coal power and hydropower, nuclear power is an E-admissible option. Following the extended weighted average principle, a permissible compromise to William and Hillary's conflict is the Bayesian value function constructed from William's probability assessment and Hillary's evaluations of the possible consequences. According to this value function, nuclear power is optimal.⁴⁴

This difference in admissibility illustrates that while the situations in figures 5.9 and 5.11 may give rise to identical orderings—and even identical numerical evaluations—of the options by William and Hillary respectively, the different sorts of reasons underlying the evaluations in each situation involve different areas of conflict. Consequently, Levi recommends different sorts of compromises for each situation. When William and Hillary's conflict is limited to probability distributions, nuclear power is inadmissible, but when they also conflict over their evaluations of the consequences, nuclear power becomes admissible. So according to this account of admissibility it is not enough to know the rival orderings of the options, or even the rival value functions behind these orderings should they exist, to determine whether an option is admissible according to this framework. In this way, Levi seeks to separate out all the reasons behind a conflict, determine all possible compromises—using weighted averaging whenever possible—for each of the conflicting reasons, and then reassemble these compromises into a more general set of compromises concerning the available options.

Similar to maximality, there might be a concern that admissibility is not strict enough, for in cases of severe conflict, scanning all the possible compromises might reveal that all the available options are also admissible. Once again, however, if

⁴³My comment at footnote 39 in this chapter applies here as well.

⁴⁴Denoting this function as v , $v(\text{nuclear}) = 1.1$, $v(\text{coal}) = 0.6$, and $v(\text{hydro}) = 0.6$. Thus $v(\text{nuclear}) > v(\text{coal})$ and $v(\text{nuclear}) > v(\text{hydro})$.

	$\mathcal{V}_1$		$\mathcal{V}_2$	
	Job Creation	Maintenance Costs	Environmental Impact	Building Costs
coal	0.0	1.0	0.0	0.9
nuclear	0.1	0.1	1.0	0.0
hydro	1.0	0.0	0.9	1.0

FIGURE 5.12: Four criteria for ranking coal power, nuclear power, and hydropower are separated into two profiles of von Neumann-Morgenstern value functions $\mathcal{V}_1$ and $\mathcal{V}_2$, where profile $\mathcal{V}_1$ having lexical priority over $\mathcal{V}_2$.

only one option may ultimately be implemented but several V- or E-admissible options are returned, then second-tier considerations may further whittle the admissible options. For instance, consider the two profiles of value functions in figure 5.12. According to the value functions in $\mathcal{V}_1$, coal power and hydropower are $\mathcal{V}_1$ -admissible while nuclear power is eliminated as second worst. If it is only possible to implement a single option, then the value functions in $\mathcal{V}_2$ identify hydropower as the only $\mathcal{V}_2$ -admissible option from coal power and hydropower. As noted when doing a similar lexical process with categorical maximality, such a tiered structure of value functions requires agreement that those in $\mathcal{V}_1$ have very strict priority over those in $\mathcal{V}_2$. In addition, such a lexical rule for admissibility, marks further divergence from revealed preference theories of choice.⁴⁵

5.2.1 Some Difficulties With Admissibility

In general, I believe that Levi's account of admissibility is an improvement over categorical maximality for two reasons. First, in addition to considering the options that are optimal according to the initial rival value judgments, it looks to options that are also optimal according to the potential resolutions and compromises between these judgments. This explicitly encourages the decision-making process to seek out potential compromises as opposed to exclusively considering the initial judgments. Categorical maximality only does the latter. Second,

⁴⁵In addition to violating Sen's properties β and γ (see footnote 36 in this chapter) the lexical V-admissibility method of choice violates Sen's property α , Sen, 'Choice Functions and Revealed Preference', p. 313. According to this property, if A is a subset of options in B , option x is in both, and x is chosen from B , then x is chosen from A as well. Put informally, a winner (x) cannot become a loser by taking away other available options (going from B to A). To see how lexical V-admissibility violates this, suppose that $\mathcal{V}_1$ is as given in figure 5.12 while $\mathcal{V}_2'$ only has the 'environmental impact' function from that figure. When given all three options, $\mathcal{V}_1$ eliminates nuclear power, $\mathcal{V}_2'$ then eliminates coal power, and hydropower is the only admissible option. However, when given just hydropower and nuclear power, both are $\mathcal{V}_1$ -admissible but $\mathcal{V}_2'$ eliminates hydropower. So hydropower loses is no longer a 'winner' when an other option, coal power, is taken away, thus violating property α .

admissibility explicitly allows for inter- or intra-personal and cardinal considerations whenever they are available, as relevant for determining these resolutions. Categorical maximality only begins to take such information into account when it incorporates judgments concerning neutral probability mixtures. Some concerns emerge, though, with Levi's reliance on using weighted averaging as the sole determinant of a resolution. Weighted averaging does have the benefit that it provides a simple procedure for determining resolutions and compromises, but the problem is that it may exclude some potential resolutions while also including some that are unacceptable.

The problem with the exclusivity of weighted averaging is familiar from criticisms of sum ranking in classical utilitarianism.⁴⁶ For instance, considerations relying on maximin or strict sufficiency are not expressible in terms of weighted averages. According to maximin, the value of an option is solely derived from the worst evaluation it receives across all the rival value judgments. So a profile of value functions $\mathcal{V}$ may be used to construct a value function according to the following rule:

Maximin: $v_{\text{Min}}(x) = \min_{i=1}^n [v_i(x)]$, for each option x .⁴⁷

This function specifies the 'min' component of maximin; selecting the optimal option identified by v_{Min} provides the 'max' component. So in the case of profile $\mathcal{V}_*$ in figure 5.5, only nuclear power is optimal according to v_{Min} , and, as figure 5.7 shows, there is no weighted average also doing this.⁴⁸ As a result, maximin reveals a possible resolution to the conflict that is ignored by the weighted average principle.⁴⁹

⁴⁶I should add that this is not to say that the weighted average principle is identical with classical utilitarianism. The former considers all possible weighted averages of the value judgements, while the latter only considers the one where they are each given equal weight. Levi, 'Amartya Sen', pp. 65–66, emphasizes this claim.

⁴⁷Itzhak Gilboa and David Schmeidler, 'Maximin Expected Utility with Non-Unique Prior', *Journal of Mathematical Economics* 18, 2 (April 1989): 141–153, presents the Γ -maximin rule, which applies maximin reasoning to instances of involving uncertainty.

⁴⁸Applying v_{Min} here shows that $v_{\text{Min}}(\text{coal}) = 0$, $v_{\text{Min}}(\text{nuclear}) = 0.1$, and $v_{\text{Min}}(\text{hydro}) = 0$. There is no single weighted average of the functions in $\mathcal{V}_*$ that is consistent with such evaluations.

⁴⁹Maximin considerations are most famously applied to value functions measuring individual holdings of primary goods in John Rawls, *A Theory of Justice* (Cambridge, MA: Belknap Press of Harvard University Press, 1971), pp. 152–156. Levi does not ignore this type of reasoning, but he argues that it is best understood as a second-tier, security consideration, Levi, *Hard Choices*, pp. 113–114, 173–174. To illustrate, suppose that $\mathcal{V}_1$ consists of the value functions from profile $\mathcal{V}_*$ in figure 5.5. In that case, only coal power and hydropower are $\mathcal{V}_1$ -admissible. Using the same set of value functions, a v_{Min} may be constructed. If this function alone comprises $\mathcal{V}_2$ then both coal power and hydropower are $\mathcal{V}_2$ -admissible. According to Levi, a second worst option is just too poor of an option to choose, even if it is the most 'secure' option. Even so, notice that had $\mathcal{V}_1$

A strict sufficiency approach, on the other hand, identifies options that surpass some pre-specified threshold.⁵⁰ That is, let $\Theta = \langle \theta_1, \theta_2, \dots, \theta_n \rangle$ be the collection of (possibly unavailable) options identifying the threshold for each value function in $\mathcal{V}$, where θ_i is an option just meeting the sufficiency threshold for v_i . According to strict sufficiency, an option is only judged on whether it meets these thresholds for $\mathcal{V}$:

Strict Sufficiency:

$$v_{\Theta}(z) = \begin{cases} 1 & \text{if for all } v_i \in \mathcal{V}, v_i(z) \geq v_i(\theta_i), \\ 0 & \text{otherwise.} \end{cases}$$

Suppose thresholds for the functions in profile $\mathcal{V}_*$ are each set at 0.05. In that case, only nuclear power is optimal according to v_{Θ} .⁵¹ As in the case of maximin, there is no weighted average of the value functions in $\mathcal{V}_*$ consistent with such a judgment.⁵² So both maximin and strict sufficiency considerations are overlooked when adhering solely to the weighted average principle.

Using weighted averages also runs the risk of being too inclusive concerning what should count as a permissible resolution to conflicting value orderings. To start with, Teddy Seidenfeld, Joseph Kadane, and Mark Schervish present two problems with the extended average principle. The first problem is that this principle may consider compromises that violate the weak Pareto principle when there is conflict concerning both the probability distributions and the evaluations of the possible consequences.⁵³ That is, two value functions may agree that one option is superior to another, and yet the extended weighted average principle generates

instead consisted of the value functions from profile $\mathcal{V}^*$ in figure 5.5, nuclear power would have been V_1 -admissible and then the uniquely V_2 -admissible option. So when an option is second best, Levi's security considerations can favor it.

⁵⁰Sufficiency approaches to decision making in political contexts are defended by Jeremy Waldron, 'John Rawls and the Social Minimum', *Journal of Applied Philosophy* 3, 1 (March 1986): 21–33; Harry Frankfurt, 'Equality as a Moral Ideal', *Ethics* 98, 1 (October 1987): 21–43; and Elizabeth S. Anderson, 'What Is the Point of Equality?', *Ethics* 109, 2 (January 1999): 287–337.

⁵¹If William's function is v_1 and Hillary's is v_2 , then $v_1(\theta_1) = v_2(\theta_2) = 0.05$. Consequently, $v_{\Theta}(\text{coal}) = 0$ because $v_2(\text{coal}) < v_2(\theta_2)$, and $v_{\Theta}(\text{hydro}) = 0$ because $v_1(\text{hydro}) < v_1(\theta_1)$. $v_{\Theta}(\text{nuclear}) = 1$, however, because $v_1(\text{nuclear}) > v_1(\theta_1)$ and $v_2(\text{nuclear}) > v_2(\theta_2)$.

⁵²This is very similar to the argument made in Sen, 'Incompleteness and Reasoned Choice', p. 52. In response to this, Levi, 'Amartya Sen', pp. 63–64, argues that since the functions in $\mathcal{V}$ are supposed to be von Neumann-Morgenstern value functions, above threshold option (like nuclear power in my example) should be V -admissible as well. So the weighted average principle can capture threshold concerns even when there is no single weighted average consistent with v_{Θ} .

⁵³Teddy Seidenfeld, Joseph B. Kadane and Mark J. Schervish, 'On the Shared Preferences of Two Bayesian Decision Makers', *Journal of Philosophy* 86, 5 (May 1989): 225–244, pp. 238–241. This arises as a consequence of their impossibility result, which was discussed in section 3.3. Recall

	Dick	Jane
r^*	1.0	1.0
r	0.1	0.4
r_*	0.0	0.0

FIGURE 5.13: *Dick and Jane's respective evaluations of the three possible consequences, r^* , r , r_* , for treating a patient.*

compromises that do not respect this simple sort of agreement. To slightly modify the example presented by Seidenfeld et al., suppose that Dick and Jane are doctors choosing between treatments T_1 and T_2 for a patient.⁵⁴ For either treatment, there are three possible consequences to the patient, (r_*) death, (r) recovery with chronic side effects, and (r^*) full recovery without side effects. Figure 5.13 shows how Dick and Jane evaluate these consequences. So both doctors agree on their evaluations concerning death and full recovery, but disagree over where recovery with side effects lies between them. Dick maintains that such side effects would pose a serious blow to the patient's quality of life, while Jane believes these side effects are serious but the patient should still be able to function fairly well.

Now the effectiveness of each treatment depends on whether the patient is allergic to treatment T_1 , and Dick and Jane disagree about this as well. Dick believes that there is only a 10 percent chance that the patient is allergic whereas Jane believes that it is a bit more likely that the patient is allergic, putting it at 30 percent. Extensive studies accepted by both doctors, however, reveal the probable consequences for administering each treatment given the patient's allergenic state, which is presented in the decision matrix in figure 5.14. To explain, if the patient is allergic, then administering T_1 carries a 78.5 percent chance of death and a 21.5 percent chance of full recovery, while administering T_2 carries a 19 percent chance of death, a 1 percent chance of full recovery, and an 80 percent chance of recovery with side effects. On the other hand, if the patient is not allergic, then administering T_1 carries a 1 percent chance of death, a 79 percent chance of full recovery, and a 20 percent chance of recovery with side effects, while administering T_2 carries a 16.5 percent chance of death and an 83.5 percent chance of full recovery.

Given all this information, both Dick and Jane agree that treatment T_2 is

that the rub of this theorem is that any Bayesian pooling function satisfying non-dictatorship ($\mathcal{D}_B$) cannot also satisfy the weak Pareto principle ($\mathcal{P}_B$). This means that given Bayesian value functions v_1 , derived from $(Pr_1, \bar{v}_1)$, and v_2 derived from $(Pr_2, \bar{v}_2)$, if $Pr_1 \neq Pr_2$ and $\bar{v}_1 \neq \bar{v}_2$, then there are situations where there is no Bayesian value function v differing from v_1 and v_2 that preserves all agreements between v_1 and v_2 concerning the superiority of one option to another.

⁵⁴ibid., pp. 230–235.

	ω_1	ω_2
T_1	$0.785(r_*) \oplus 0.215(r^*)$	$0.010(r_*) \oplus 0.200(r) \oplus 0.790(r^*)$
T_2	$0.190(r_*) \oplus 0.800(r) \oplus 0.010(r^*)$	$0.165(r_*) \oplus 0.835(r^*)$

FIGURE 5.14: *A decision matrix representing the uncertainties involved in giving either treatment T_1 or T_2 to a patient. ω_1 is the state of affairs where the patient is allergic to T_1 , and ω_2 is that state of affairs where the patient is not allergic to it. [This figure is adapted from a similar one in Seidenfeld, Kadane, and Schervish, 'Shared Preferences of Two Decision Makers', p. 233.]*

superior to treatment T_1 .⁵⁵ The weak Pareto principle therefore suggests that this agreement must be respected: any potential resolution to Dick and Jane's conflict should agree that T_2 is superior to T_1 . The extended weighted average principle, however, considers compromises that do not respect this agreement. This is because this latter principle first seeks all possible compromises by taking all weighted averages of the probabilities (so anything putting the probability of the patient being allergic within the interval 0.1 to 0.3), and then does the same for the evaluation of consequences (so anything putting the value of recovery with side effects within the interval 0.1 to 0.4 while giving death 0.0 and full recovery 1.0). Then all the possible permutations combining these two 'lower-level' compromises uniquely determine all the possible 'higher-level' compromises concerning which treatment to administer. One of these latter compromises includes combining Dick's assessment of the consequences with Jane's assessment of the probabilities, and this compromise maintains that T_1 is superior to T_2 .⁵⁶ As a result, T_1 is E-admissible, even though both Dick and Jane unanimously agree that T_2 is superior to it. Along with Seidenfeld et al., I find it unsettling that there are situations where the extended weighted average principle considers compromises that overturn such a fundamental agreement. This points to a serious problem: considering the potential resolutions and compromises with respect to the reasons underlying the rival value orderings—in this case, the probability judgments and assessment of consequences—permits admissible options to include alternatives that those orderings agree are inferior to another option. I return to this issue in section 5.3.

In a different article, Seidenfeld et al. also point out a second problem with the extended average principle: it may consider compromises that reject agreed upon assumptions concerning the probability distributions.⁵⁷ The example they offer illustrating this is a modification of the previous one. Consider the decision faced

⁵⁵According to Dick, $v_1(T_1) = 0.7505$ and $v_1(T_2) = 0.7605$, and according to Jane, $v_2(T_1) = 0.6735$ and $v_2(T_2) = 0.6835$. So both agree that $v_i(T_2) > v_i(T_1)$.

⁵⁶According to this compromise, $v(T_1) = 0.6315$ and $v(T_2) = 0.6115$. Thus $v(T_1) > v(T_2)$.

⁵⁷Seidenfeld, Schervish and Kadane, 'Coherent Choice Functions Under Uncertainty', pp. 168–169.

by Meredith and Christina, two doctors who must choose whether they should administer treatment T_3 or treatment T_4 to a patient. Analogous to the previous example, the effectiveness of either treatment depends upon whether the patient is allergic to T_3 . In this case, however, there are only two possible consequences, (r_*) death and (r^*) full recovery, and Meredith and Christina agree in their evaluations of these ($\bar{v}(r_*) = 0.00$ and $\bar{v}(r^*) = 1.00$). Both doctors also agree on how the effectiveness of each treatment given the patient's allergenic state: if the patient is allergic, then the consequence of administering T_3 is death and the consequence of administering T_4 is full recovery; if the patient is not allergic, though, T_3 leads to full recovery and T_4 leads to death. These consequences all come with 100 percent certainty. The problem in this situation is that Meredith and Christina disagree over the probability that the patient is allergic to T_3 . Meredith gives this a 40 percent chance while Christina gives it a 60 percent chance.

Now there also exists a third 'treatment' option T_5 , however, that would have Meredith and Christina spend time determining whether it is currently sunny in Paris, France, a city far away from their current location. According to T_5 , if it is sunny in Paris, the doctors administer T_3 ; if not, they administer T_4 . However the consequences for doing either treatment as a result of choosing T_5 have slightly reduced value (by 0.01 each) because of the time spent finding a weather report. As a result, there are four states of affairs relevant to this case:

- State ω_1 : The patient is allergic and it is sunny in Paris,
- State ω_2 : The patient is allergic and it is not sunny,
- State ω_3 : The patient is not allergic and it sunny, and
- State ω_4 : The patient is not allergic and it not sunny.

Both doctors agree that the associate values of the three treatments in each of these four states make up the decision matrix presented in figure 5.15. However, they disagree over the likelihood that it is currently sunny in Paris. Meredith believes there is a 20 percent chance it is sunny and Christina believes this has an 80 chance. Even so, both agree that weather conditions in Paris are medically irrelevant for determining the patient's allergic condition. Meredith and Christina's respective probability distributions over these four states of affairs appear in figure 5.16, where Pr_1 represents Meredith's distribution and Pr_2 represents Christina's.

Given this information and a choice from the three treatments, both T_3 and T_4 are E-admissible because T_3 is optimal according to Meredith and T_4 is optimal according to Christina.⁵⁸ The concern is that T_5 is also E-admissible. This is because the extended weighted average principle considers any weighted average

⁵⁸According to Meredith, $v_1(T_3) = 0.60$, $v_1(T_4) = 0.40$, and $v_1(T_5) = 0.55$. According to Christina, $v_2(T_3) = 0.40$, $v_2(T_4) = 0.60$, and $v_2(T_5) = 0.55$.

	ω_1	ω_2	ω_3	ω_4
T_3	0.00	0.00	1.00	1.00
T_4	1.00	1.00	0.00	0.00
T_5	0.99	-0.01	-0.01	0.99

FIGURE 5.15: A decision matrix representing the uncertainties involved in giving either treatment T_3 , T_4 , or T_5 to a patient. [This figure is adapted from a similar one in Seidenfeld, Schervish, and Kadane, 'Coherent Choice Functions Under Uncertainty', p. 168.]

	ω_1	ω_2	ω_3	ω_4
Pr_1	0.08	0.32	0.12	0.48
Pr_2	0.48	0.12	0.32	0.08
Pr_3	0.28	0.22	0.22	0.28

FIGURE 5.16: Three different probability distributions for the four potential states of affairs affecting Meredith and Christina's decision. [This figure is adapted from a similar one in Seidenfeld, Schervish, and Kadane, 'Coherent Choice Functions Under Uncertainty', p. 169.]

of Meredith and Christina's probability distributions as designating a permissible compromise for their conflict. Unlike the previous case, there is no need here to consider weighted averages of the evaluations of consequences since there is no disagreement concerning them. So one such weighted average of probabilities is where Meredith's distribution receives half the weight and Christina's the other half, which is the distribution Pr_3 from figure 5.16. Given this distribution along with the values from the decision matrix, T_5 is superior to both T_3 and T_4 and therefore E-admissible.⁵⁹ This is problematic because T_5 is an option that both Meredith and Christina presumably agree is unacceptable because it involves paying to discover medically irrelevant meteorological conditions in order to determine which treatment to administer.⁶⁰ This suggests the deeper problem: Pr_3 may make T_5 E-admissible, but this distribution expressly denies the probabilistic independence of the patient's allergic state from the weather in Paris.⁶¹ According

⁵⁹According to this value function, $v(T_3) = 0.50$, $v(T_4) = 0.50$, and $v(T_5) = 0.55$.

⁶⁰Of course, it would be quite reasonable for the doctors to pay to discover, for instance, medically relevant genetic information concerning the patient that could influence the probability of the patient being allergic.

⁶¹Two events A and B are probabilistically independent if and only if $\text{Pr}(A) = \text{Pr}(A|B)$. That is, the probability of event A is not affected by whether B also occurs or not. In this case, if A represents the event that the patient is allergic and B represents that it is sunny in Paris, both Pr_1 and Pr_2 from figure 5.16 do reflect the fact that both Meredith and Christina regard A and B as independent events. This is not true for Pr_3 because $\text{Pr}_3(A) = 0.50$ while $\text{Pr}_3(A|B) = 0.56$.

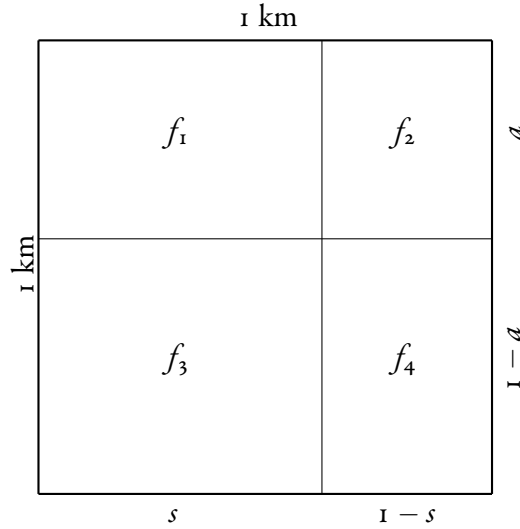


FIGURE 5.17: *A general schematic of the four growing fields in Mario and Luigi's estate. The precise locations of the two interior walls separating the fields are uncertain.*

to Prob₃, the two events are positively correlated, which both doctors agree is false. So once again, the extended weighted average principle is too inclusive. In this case it requires considering probability distributions that may violate a fundamental agreement between the original probability distributions concerning the relationship between possible events.

In fact, I argue that a similar problem with the weighted average principle for V-admissibility emerges, even when there is no disagreement concerning probabilities. For instance, suppose two estranged brothers, Mario and Luigi, live in the United States, and recently they jointly inherited their father's small estate in Italy. According to the father's will, both brothers must agree upon any decision concerning the estate. Each brother sends his personal assayer to report on the estate and both reports agree that it consists of a single plot of land that is one kilometer by one kilometer in area. Both reports also agree that this area is divided into four rectangular fields by exactly two straight one-kilometer walls. These claims about the field are depicted in figure 5.17.

The brothers' reports also note that only two of the four fields may be cultivated at the same time for a given season while the other two must lie fallow. Based on these reports, both brothers agree that any value function assessing the profitability

	f_1	f_2	f_3	f_4
A_1	fallow	fallow	wheat	wheat
A_2	wheat	wheat	fallow	fallow
A_3	corn	fallow	fallow	corn
A_4	wheat	fallow	fallow	wheat
A_5	fallow	wheat	wheat	fallow

FIGURE 5.18: Five different crop allocations for the fields on Mario and Luigi's estate.

of a distribution of crops across the four fields should have the following form:

$$\begin{aligned}\hat{v}(x) = & (a \times s \times f_1(x)) + (a \times (1 - s) \times f_2(x)) \\ & + ((1 - a) \times s \times f_3(x)) + ((1 - a) \times (1 - s) \times f_4(x)).\end{aligned}$$

Here f_1, f_2, f_3 , and f_4 represent the profitability of each of the four fields, in dollars per square kilometer, given the type of crop planted in it. In particular, the report notes that if wheat is planted in a field, it will yield a profit of 10.0 thousand dollars per square kilometer, whereas if corn is planted in an area, it will yield a profit of 9.8 thousand dollars per square kilometer.

The two reports, however, disagree about the precise location of those two dividing walls. According to Mario's report, the length of a is 0.4 kilometers and the length of s is 0.2 kilometers. So as far as he knows, field f_2 , for instance, encloses an area of 0.32 square kilometers. Meanwhile, Luigi's assayer claims that the length of a is 0.6 kilometers and the length of s is 0.8 kilometers. So according to this assessment, field f_2 is 0.12 square kilometers in size. As a result, Mario and Luigi have competing value functions concerning the profitability of a season's crops. Mario's value function is

$$v_1(x) = (0.08 \times f_1(x)) + (0.32 \times f_2(x)) + (0.12 \times f_3(x)) + (0.48 \times f_4(x)).$$

Meanwhile, Luigi's function is different:

$$v_2(x) = (0.48 \times f_1(x)) + (0.12 \times f_2(x)) + (0.32 \times f_3(x)) + (0.08 \times f_4(x)).$$

Both brothers are dismayed at this divergence, but before they can send their assayers back to recheck their measurements, the brothers are told that they must immediately decide what to plant in two of the four fields. They are given the first three crop allocations— A_1, A_2 , and A_3 —from figure 5.18.

Based on his value function, Mario maintains that allocation A_1 is optimal, while Luigi disagrees, because according to his value function A_2 is optimal.⁶² So

⁶²According to Mario, $v_1(A_1) = \$6,000$, $v_1(A_2) = \$4,000$, and $v_1(A_3) = \$5,488$. According to Luigi, $v_2(A_1) = \$4,000$, $v_2(A_2) = \$6,000$, and $v_2(A_3) = \$5,488$.

both A_1 and A_2 are V-admissible. Furthermore, allocation A_3 is also V-admissible because it is optimal according to the 50-50 weighted average of v_1 and v_2 :

$$\begin{aligned} v_{50-50}(x) &= (0.5 \times v_1(x)) + (0.5 \times v_2(x)) \\ &= (0.28 \times f_1(x)) + (0.22 \times f_2(x)) + (0.22 \times f_3(x)) + (0.28 \times f_4(x)).^{63} \end{aligned}$$

The problem with this, however, is that v_{50-50} does not have the form that the two brothers agreed upon. That is, Mario and Luigi maintain that the permissible value functions representing their situation should have the form given in $\hat{v}$ above, where the weights on the f_i 's can be expressed in terms of s and a . It turns out that v_{50-50} simply cannot do this.⁶⁴ Once again, taking weighted averages considers resolutions that are unanimously rejected by the original rival value functions.

Given Levi's approach as I have outlined it in this section, I suspect that Levi would recommend that the brothers do not apply the weighted average principle to their rival value functions v_1 and v_2 . Instead, Levi might pose using something like the extended weighted average principle, having Mario and Luigi distinguish the conflicting components making up these value functions. Since the root of this conflict concerns the values of a and s , all the possible weighted averages of each value could be determined separately, and then a compromise constructed by combining one weighted average of the conflicting a 's and one weighted average of the conflicting s 's. As a result, the resulting value function would then have the form of $\hat{v}$ above, with an a ranging from 0.4 to 0.6 and an s ranging from 0.2 to 0.8. One such combination combines Mario's value for s and Luigi's value for a , resulting in the function:

$$v_3(x) = (0.12 \times f_1(x)) + (0.48 \times f_2(x)) + (0.08 \times f_3(x)) + (0.32 \times f_4(x)).$$

Unfortunately, this suggestion suffers from the same problem faced by Dick and Jane's use of the extended weighted average principle. Consider crop allocations A_4 and A_5 from figure 5.18. Both Mario and Luigi agree that A_4 is superior to A_5 , but the compromise v claims instead that A_5 is superior to A_4 .⁶⁵ So in a choice between only A_4 and A_5 both are admissible even though both parties agree that A_5 should not be chosen when A_4 is available. Once again, focusing on potential

⁶³ According to this weighted average, $v_{50-50}(A_1) = v_{50-50}(A_2) = \$5,000$, and $v_{50-50}(A_3) = \$5,488$.

⁶⁴ In order for v_{50-50} to have the same form as $\hat{v}$, then the following conditions must all hold: (1) $a \times s = 0.28$, (2) $a - (a \times s) = 0.22$, (3) $s - (a \times s) = 0.22$, and (4) $1 - a - s + (a \times s) = 0.28$. For (1) and (2) to both hold, $a = 0.50$ and $s = 0.56$. But if that is the case, then $s - (a \times s) = 0.28$, which contradicts (3), and $1 - a - s + (a \times s) = 0.22$, which contradicts (4). So there is no way to assign values for a and s satisfying conditions (1)–(4).

⁶⁵ According to both Mario and Luigi, $v_i(A_4) = \$5,600$ and $v_i(A_5) = \$4,400$, but according to v , $v(A_4) = \$4,400$ and $v(A_5) = \$5,600$.

compromises concerning the reasons underlying the brothers' orderings—now the two dimensions of the field—results in a violation of the weak Pareto condition.

Together, all these difficulties suggest that relying solely on the (extended) weighted average principle to delineate potential resolutions and compromises is too imprecise. This principle ignores potential resolutions such as maximin or strict sufficiency as well as some potentially important agreements, such as those concerning the inferiority of one option to another, probabilistic independence of events, or even more general formal considerations. In light of all these problems, I maintain that permissible resolutions and compromises, which I do believe should still be consulted, must be responsive to what the decision makers consider to be the relevant information contained in the rival value orderings. However, I do not believe that there is any mechanical process—weighted averaging or otherwise—for anticipating in advance what counts as the 'relevant' information. Such information is often sensitive to the issue at hand and the perceived interests of those involved or effected by the decision-making process. Before going on to address what this means for liberal democracy, however, I must first explain why I maintain, contrary to Levi, that decision making under unresolved conflict should only consider potential resolutions and compromises between rival assessments of options and *not* between the reasons underlying those assessments.

5.3 Agreement on Reasons Versus Agreement on Options

The example of Mario and Luigi from the previous section sets up a dilemma for Isaac Levi's claim that potential agreements between rival value assessments ought to be understood in terms of weighted averaging. This is because Levi's framework is extremely sensitive to the difference between (1) reaching an agreement on the reasons for selecting an option and (2) reaching an agreement on the options themselves.⁶⁶ If the focus is on the first sort of agreement, one should examine the reasons behind rival value judgments and consider all the potential agreements between those reasons that conflict. These agreements over reasons then determine the permissible ways in which to evaluate the options. For Mario and Luigi, v_1 and v_2 conflict because there is a disagreement concerning the values of s and a , and so any combination of their respective weighted averages generates an permissible agreement. This preserves the requisite form $\hat{v}$ of an evaluation, which Mario and Luigi desire, but then violates weak Pareto by declaring A_5 admissible when A_4 is the only other available option. In short, seeking agreements on the reasons

⁶⁶See, for instance, Seidenfeld, Kadane and Schervish, 'Shared Preferences of Two Decision Makers', p. 241.

via weighted averaging overturns a fundamental agreement on options. Should Mario and Luigi instead focus on the second sort of agreement, agreement on the options themselves, all weighted averages of v_1 and v_2 are permissible. This respects weak Pareto, but these weighted averages include assessments like v_{50-50} which do not possess the requisite form $\hat{v}$. So even directly seeking agreements on the options via weighted averaging may still violate basic agreements other than those entailed by weak Pareto.

Faced with this dilemma for weighted averaging, Levi willingly falls on its first horn, arguing for a focus on agreements concerning reasons, even though doing so at times may violate the weak Pareto principle.⁶⁷ For example, I suspect that Levi's framework claims that Mario and Luigi should regard A_5 as admissible even when A_4 is the only other available option. According to Levi, however, this judgment still preserves an agreement on options, though one of a different sort than weak Pareto. That is, the weak Pareto principle demands respecting all unanimous agreements that may exist between Mario and Luigi's respective orderings of the options. Levi believes that this is too restrictive a form of Pareto. Instead, he recommends what he calls the robust Pareto unanimity principle. According to this principle, it is agreement between Mario, Luigi, and *two additional assessments* that must be respected. The first of these additional assessments comes from a hypothetical Maria who accepts Mario's value for s but Luigi's value for a (this is value function v_3 —the one judging A_5 as optimal—from the previous section). The other assessment comes from a hypothetical Luisa who accepts Luigi's value for s but Mario's value for a . It turns out that considering agreements on Mario and Luigi's reasons does respect unanimous agreements concerning how Mario, Luigi, Maria, and Luisa assess the options.⁶⁸ Something like weak Pareto remains, but in some instances it may require postulating non-existent parties like Maria and Luisa.

I find Levi's position untenable for two reasons. First of all, it is strange that the decisions of a group of people should be responsive to the assessments of entirely hypothetical people.⁶⁹ If all the group's real members agree that one option is superior to another, I see no reason to overturn that consensus because it is possible to imagine people who might disagree. For instance, if everyone agrees that a nuclear power plant should be built instead of a coal one, then it seems reasonable to eliminate coal power as an option, even if some hypothetical person

⁶⁷Levi, 'Pareto Unanimity and Consensus', 'Value Commitments, Value Conflict', and 'Why Indeterminate Probability is Rational', *Journal of Applied Logic* 7, 4 (December 2009): 364–376.

⁶⁸Seidenfeld, Kadane and Schervish, 'Shared Preferences of Two Decision Makers', pp. 241–242; and Levi, 'Pareto Unanimity and Consensus', pp. 489–492.

⁶⁹Levi's respect for the evaluations of hypothetical individuals should not be confused with Rawls' strategy with the original position. As I discussed in section 2.3.1.... **Finish this later.**

might have reason to prefer it to nuclear power. Certainly if such a contrary person eventually joins the group, or if one of the existing members changes her mind on the matter, it may be necessary to reconsider or even revise that judgment. Even so, without actual people endorsing those views, however, it seems problematic to give veto power over a decision to entirely hypothetical persons. There seems little legitimacy for the state to make a decision that all those actually subject to it would unanimously reject in favor of another alternative.

This problem may not be too devastating for Levi's approach, however, because I believe a potential solution has already been suggested. The concern with considering hypothetical persons or considering agreements on reasons via weighted averaging, both of which amount to the same thing, is that doing so may be too liberal in identifying admissible options. Mario and Luigi, for instance, should regard A_5 as admissible because it is optimal according to the hypothetical Maria. Similarly, Dick and Jane should regard T_1 as admissible because it is optimal according to a hypothetical Mary, who has Dick's assessment of the consequences and Jane's assessment of the probabilities. Perhaps it is more reasonable to consider the views of hypothetical individuals when making an initial determination of admissibility, but then eliminate any options that are unanimously agreed to be inferior to another option by the actual individuals. More generally, this suggests a lexical process: (extended) weighted averaging determines first-tier admissibility but second-tier criteria such as weak Pareto further winnow down the admissible alternatives.⁷⁰ So A_5 for Mario and Luigi, or T_1 for Dick and Jane, might make the first cut, but then be eliminated by standard weak Paretian considerations. The drawback with this lexical approach is that it not clear what purpose weighted averaging serves in the first-tier when second-tier considerations subsequently eliminate the potential compromises it identifies.⁷¹

There is a second, more general, problem with the focus on securing agreements on reasons. This concerns the possibility of regress when attempting to isolate the fundamental reasons of a conflict. I can make this problem clearer by first showing how it arises in a different context with the doctrinal paradox.⁷² According to this

⁷⁰Put this generally, a similar lexical process could potentially resolve the concern that Meredith and Christina must regard T_5 as E-admissible. In this case, second-tier considerations involving probabilistic independence could eliminate this option.

⁷¹Furthermore, this solution does not avoid the problem of weighted averaging already ignoring potential compromises in the form of maximin and strict sufficiency.

⁷²Lewis A. Kornhauser and Lawrence G. Sage, 'The One and the Many: Adjudication in Collegial Courts', *California Law Review* 81, 1 (January 1993): 1–59, p. 11. This paradox can be reformulated as an impossibility theorem along the lines of Arrow's theorem, see Christian List and Philip Pettit, 'Aggregating Sets of Judgments: An Impossibility Result', *Economics and Philosophy* 18, 1 (April 2002): 89–110, and 'Aggregating Sets of Judgments: Two Impossibility Results', *Synthese* 140, 1–2 (May 2004): 207–235, pp. 212–215.

	Was the crime willful?	Was the crime premeditated?	Impose the death penalty?
Juror 1	Yes	No	No
Juror 2	Yes	No	No
Juror 3	Yes	Yes	Yes
Juror 4	No	Yes	No
Juror 5	No	Yes	No

FIGURE 5.19: *Five jurors and their respective positions concerning a defendant's crime and whether it warrants the death penalty.*

paradox, suppose a man is found guilty of a crime, which carries with it the possible punishment of execution dependent upon certain circumstances. In particular, the law states that the defendant should receive the death penalty for this crime if and only if the crime is deemed both willful and premeditated. In this case, a jury of five individuals has the views as depicted in figure 5.19. There, one 3-to-2 majority of jurors believes the crime was willful while another 3-to-2 majority believes it was premeditated. Should the jurors then use majority rule to reach agreement on each reason for selecting a punishment, the jury rules to execute the defendant because there is a majority accepting each reason that jointly entails imposing that penalty. However, if the jurors directly use majority rule to determine whether to execute the defendant or not, 4-to-1 majority is against imposing the death penalty. Reaching agreements on the reasons for selecting the punishment leads to execution; reaching agreement on the punishment itself avoids execution, though in this case, agreement is secured by majority rule and not weighted averaging. This distinction between the two sorts of agreements has tremendous consequences for the defendant, and hence the dilemma for majority rule.

Now suppose that Juror 3, maintains that the jury should secure agreement via majority rule on the reasons for selecting an option, and so maintains that the defendant should be executed. The regress problem emerges because the disagreement over the reasons given in figure 5.19 might be the result of even deeper reasons. For instance, Juror 1 might get everyone to agree that that for the crime to have been willful, the defendant must either not have had diminished capacities or had an intent to cause harm. Furthermore, he might convince them that the crime was premeditated if either the intent was to cause harm or the crime was planned in advance. The jurors judgments concerning these even deeper reasons behind determining the crime's penalty are shown in figure 5.20. This reveals that a 3-to-2 majority believes that the defendant had diminished capacities, a 4-to-1 majority believes there was no intent to cause harm, and a 3-to-2 majority believes the crime was not planned in advance. If these majorities represent the

	Did the defendant have diminished capacities?	Was the intent to cause harm?	Was the crime planned in advance?
Juror 1	No	No	No
Juror 2	No	No	No
Juror 3	Yes	Yes	No
Juror 4	Yes	No	Yes
Juror 5	Yes	No	Yes

FIGURE 5.20: *The same five jurors from figure 5.19 and their respective positions justifying their reasoning concerning the defendant's crime presented in that figure.*

	Derek	Owen
T_A	0.0300	0.0400
T_B	0.0100	0.0100

FIGURE 5.21: *Derek and Owen's respective Bayesian value functions for assessing two possible treatments, T_A and T_B , for a patient.*

agreements on reasons, then the death penalty is no longer an option. Of course, if the goal remains to secure agreement on reasons, Juror 3 might try to dig for even deeper reasons where majority agreements would entail imposing the death penalty. In sum, when seeking to secure agreement on reasons, it remains unclear what should count as the set of privileged reasons for making the final decision. As I will argue in the next section, such a concern seems extremely difficult to overcome when presented with a reasonable pluralism of comprehensive doctrines, each of which may understand a different set of reasons as fundamental.

An example, adapted from one formulated by Teddy Seidenfeld, reveals a similar problem for Levi's focus on seeking agreements on reasons with E-admissibility when it comes to the partition of states of affairs.⁷³ Suppose two doctors, Derek and Owen, agree that treatment T_A is superior to treatment T_B , though they disagree about how much better it is, as shown by their Bayesian value functions, whose evaluations of the two treatments are depicted in figure 5.21. Given just this information, T_A is uniquely V-admissible.

Levi's framework, however, asks the doctors to explore the reasons underlying their conflicting assessments. It turns out that one root of this disagreement results from Derek and Owen assigning different probabilities to a state of affairs ω , which holds if the patient is allergic to T_A . Derek believes ω is 90 percent likely while Owen believes it is 60 percent likely. On the other hand, $\bar{\omega}$ holds

⁷³Teddy Seidenfeld, 'Outline of a Theory of Partially Ordered Preferences', *Philosophical Topics* 21, 1 (Spring 1993): 173–189, pp. 181–183. In that same article, Seidenfeld also presents a more positive proposal, which Levi might accept, for avoiding the use of partitions that are too coarse.

	Derek		Owen	
	ω	$\bar{\omega}$	ω	$\bar{\omega}$
T_A	0.0000	0.3000	0.0000	0.1000
T_B	0.0100	0.0100	0.0100	0.0100

FIGURE 5.22: *Two decision matrices representing the rival ways in which Derek and Owen evaluate the potential consequences of administering a treatment to a patient when there are two states of affairs.*

if the patient is not allergic, and so Derek and Owen respectively believe this is 10 and 40 percent likely. A second point of disagreement concerns the relative value of T_A if the patient is not allergic to it, as seen in the decision matrix from figure 5.22. This matrix does reveal, however, that Derek and Owen do agree on the value of T_A if the patient is allergic, and that they also agree that the patient's allergenic state does not effect the value of T_B . According to Levi, the doctors must then use the extended weighted average principle, which would have them consider what is admissible according to, among others, the hypothetical doctor Preston, who accepts Derek's probability distribution but Owen's assessments of the consequences from figure 5.22. According to Preston's resulting Bayesian value function, T_B is optimal.⁷⁴ Derek and Owen still agree that T_A is optimal—their overall assessments remain as given in figure 5.22—so both T_A and T_B are E-admissible.

Once again, looking at the reasons behind a conflict leads to a violation of weak Pareto, but this is not the issue that now concerns me. For suppose that Derek and Owen are curious as to why their assessments concerning the value of T_A vary when they believe that the patient is not allergic to it. Keeping in the spirit of Levi's approach, they then examine the reasons behind these conflicting judgments. In this case, it turns out that they believe that $\bar{\omega}$ breaks into two further states of affairs, $\bar{\omega}_1$ and $\bar{\omega}_2$. If $\bar{\omega}_1$ holds, the patient is not allergic to T_A but may experience extreme excitement to the central nervous system due to it; if $\bar{\omega}_2$ holds, the patient is not allergic to T_A and will not experience such a side effect. One disagreement here concerns the probability for these possibilities. Derek believes $\bar{\omega}_1$ is 4 percent likely and $\bar{\omega}_2$ is 6 percent likely, while Owen believes each is 2 percent likely. They also disagree over the value of T_A under these states of affairs as depicted in the decision matrices in figure 5.23. Here Derek believes that the patient might actually benefit from the stimulation whereas Owen does not. Regardless, Derek and Owen's overall evaluations remain as given in figure 5.22, but now Preston's evaluations have changed because the coupling of Derek's refined probability

⁷⁴Preston's Bayesian value function assigns 0.0100 to both T_A and T_B , implying Preston judges them both equally optimal.

	Derek			Owen		
	ω	$\bar{\omega}_1$	$\bar{\omega}_2$	ω	$\bar{\omega}_1$	$\bar{\omega}_2$
T_A	0.0000	0.5000	0.1666	0.0000	0.0750	0.1250
T_B	0.0100	0.0100	0.0100	0.0100	0.0100	0.0100

FIGURE 5.23: Two decision matrices representing the rival ways in which Derek and Owen evaluate the potential consequences of administering a treatment to a patient when state $\bar{\omega}$ from figure 5.22 is partitioned into two further sub-states.

distribution with Owen's refined assessments from figure 5.23 no longer regards T_B as optimal.⁷⁵ Consequently, further refinement of the reasons behind Derek and Owen's disagreement now declares that only T_A is E-admissible. Of course, exploring the disagreements between the two matrices in figure 5.23 may lead Preston to once again change his position.

The curious feature behind both the doctrinal paradox and this example of Derek and Owen is that while exploring the reasons supporting individual judgments does not change any of those judgments, the choice of what counts as the most 'basic' level of reasons can make a radical difference on the outcome of the decision making process—at least for majority rule and E-admissibility. I find this unacceptable. Certainly the exploration of the reasons behind a disagreement is important, but if such exploration only changes the minds of hypothetical participants like Preston and not the minds of those actually involved, this exploration should not arbitrarily change the outcome of the decision. I say 'arbitrarily' for two reasons. On the one hand, it is not clear what should count as the fundamental or basic reasons for an evaluation or decision. It always seems possible to dig deeper over a disputed point, for it seems that there must be an underlying reason for that disagreement. On the other hand, if the time preceding a decision is finite and limited, the level of reasons used to ultimately decide may simply be the one arrived at before time ran out. Time alone, and not the reconsideration of how the individual evaluations rank the options, may then arbitrarily decide the outcome of the decision.

In light of all these considerations, I reject Levi's insistence on seeking out potential compromises concerning the reasons behind an evaluation of options. It is sufficient to search for potential compromises on the evaluations of options themselves while remaining faithful to weak Pareto. In the case of Derek and Owen, this certainly seems obvious. Both agree that it is not even up for debate whether T_B should be administered. It is only when digging for reasons along the lines of the extended weighted average principle that an imaginary Preston

⁷⁵According to Preston's refined Bayesian value function, T_A has a value of 0.0105 and T_B has a value of 0.0100.

begs them to reconsider. Why they should listen to such an apparition appears senseless, especially since it only takes further digging to force Preston to fall back in line.

To bring together the several strands covered so far in this chapter, I have argued (1) in favor of an account of decision making that eliminates options that not optimal according to at least one permissible value ordering, (2) against Levi's account requiring that all weighted averages of two permissible value orderings should also be permissible, (3) that permissible value orderings are based upon the actual conflicting value orderings and not upon the reasons behind these orderings, (4) permissible value orderings should respect all unanimous judgments concerning the options in accord with the weak Pareto principle, and (5) that further unanimous judgments, such as those concerning probabilistic independence, should also be respected, but that the particularities of these may vary in different decision contexts. These are fairly general claims concerning decision making under unresolved conflict. Now for the remainder of this chapter, I turn to their application for decision making in the liberal state and why I believe doing so weakens the apparent tension between state neutrality and state agency that was introduced in section 1.4.

5.4 Agreement in the Liberal State

For the political decision making process, the demand to focus on reaching an agreement concerning the assessments of options, and not on agreements concerning the reasons for selecting an option, may initially seem strange and possibly contrary to the goals of liberal democracy. After all, it seems natural for liberalism to advocate that prior public deliberation concerning the exercise of state power should involve due consideration of the reasons behind the state taking that action. I have no issue with this. Deliberative democrats are correct that deliberation provides opportunities for the state's constitutive members to share the reasons each has for his or her respective position, and use these reasons to attempt to convince each other of the veracity of these positions. Sharing reasons in this manner may then cause participants to weed out pernicious thinking, change their minds, alter their assessments of the options, and perhaps even converge around accepting a single option. Furthermore, revealing the reasons behind positions through public deliberation opens up the potential for participants to discover or create new options that might secure more widespread support. Failing to deliberate on the reasons behind the individual value assessments certainly makes

it extremely unlikely for any of this to happen.⁷⁶

However, notice that this account of deliberation is primarily about changing the initial value orderings of options and creating new options to place within those rankings. My concern in the last chapter was about procedures such as majority rule and E-admissibility that do neither of these things and yet place a focus on reasons may lead to outcomes that all the state's constitutive members may agree—even after deliberation—are inferior to another option. Overturning weak Pareto, or other unanimous agreements such as those concerning probabilistic independence, in this manner seems grossly illegitimate. In addition, as I have already mentioned, it seems grossly arbitrary, especially considering how the length of deliberation may inappropriately influence the outcome. For if deliberation seeks agreement on reasons, a final outcome or decision may be extremely sensitive to the length of deliberation even when no one's position on the issues was changed and no new options were included. In the example of the five jurors, for instance, suppose deliberation lasts for time t and in that time the set of reasons given in figure 5.19 are identified, leading majority rule on reasons to demand executing the defendant. But had deliberation continued on for time $t + e$, a different set of reasons, such as those in figure 5.20 may have been identified, now causing majority rule on reasons to reject execution. The example of Derek and Owen may accordingly present similar diachronic problem for E-admissibility. In either case, it is troubling that the length of deliberation on reason may influence the outcome even when deliberation does not change any individual's position on the issue—or in the case of E-admissibility, deliberation only changes the minds of purely hypothetical individuals and not any of the actual participants.

A third problem also exists for attempting to secure agreement on reasons: participants may not even agree on the proper set of reasons to use in making this decision. In the case of the jurors, Juror 3 might favor the partition in figure 5.19 while Juror 1 favors that in figure 5.20. Similarly, Derek might understand the decision problem as given for him in figure 5.22 as relevant, whereas Owen understands it as given in figure 5.23.⁷⁷ In the first case, it is unclear which partition is appropriate for applying majority rule, and in the second case, it is unclear which applies for E-admissibility. Notice, however, that despite whichever partition is selected, the individual value assessments remain unchanged. Hence it is much more straightforward for the decision procedure to respect whatever agreements

⁷⁶These influences of deliberation are staples of most contemporary approaches to conflict resolution, the classic account being 'principled negotiation' from Roger Fisher, William Ury and Bruce Patton, *Gett to Yes: Negotiating Agreement with Giving In*, 2nd ed. (New York: Penguin Books, 1991).

⁷⁷Seidenfeld, 'Outline of a Theory of Partially Ordered Preferences', 188n14 presents a similar concern.

that may exist between these overall value judgments.

This last point seems even more pressing when recalling the discussion in section 1.4 of liberalism's desire for the state to remain neutral, or as I prefer to call it, responsively impartial, in the face of reasonable pluralism. In a liberal state, its constitutive members are permitted to endorse incompatible comprehensive doctrines—provided that these views are reasonable in the appropriate sense specified in section 1.2.2—which in turn allows these people to conflict over the reasons they find relevant and disagree over how these reasons fit together for making a decision. In the case of the city selecting which power plant to build, a green radical would almost certainly adopt a different comprehensive doctrine for assessing the options than a venture capitalist would. As a result, the reasons that the green radicals might present for relying on renewable energy at all costs might have no meaning to the venture capitalists who only consider the most profitable options. In addition, even while the green radicals may agree on the reasons they consider relevant, they might disagree amongst themselves concerning how these reasons come together when jointly assessing the options. The same may hold true for the venture capitalists. Deliberation might close these gaps, say when the radical greens respond to the concerns of the venture capitalists by adopting arguments that attempt to demonstrate the long-term profitability of renewable energy, but such changes are not guaranteed.

Even so, the liberal state's adherence to impartiality suggests that it cannot attempt to secure consensus on which reasons are appropriate for making the decisions, for in doing so, the state may have to take controversial stands on what reasons are relevant and how these should come together to form overall judgments concerning the options. So, for example, if neither the green radicals nor the venture capitalists will budge on their respective comprehensive doctrines, the state may have to impose what it takes to be the correct reasons for making the decision. This raises the concern that the liberal state cannot be responsively impartial, but may have to promote a comprehensive doctrine of its own.

Finally, and perhaps most importantly, I maintain that it is simply mistaken to assert that a person's reasons for judging options are wholly separable from each other as well as from that person's direct assessment of these options as Levi's approach suggests. The purpose of a comprehensive doctrine is to integrate reasons and overall assessments together into a coherent whole. Consequently, altering one reason might not only change a person's judgment of the options, but also change the other reasons she considers relevant, thus leading to an entirely new comprehensive doctrine. That is to say, should a radical green begin to couch her argument for renewable energy in terms of profitability, this may lead her to supplant her direct concern with the environment with other concerns such as the job creation afforded by a changeover to green technology. In sum, she

may now accept an entirely different comprehensive doctrine, one embodied by eco-capitalism, while her former colleagues among the green radicals see her as an alien, a sellout.

Therefore, I believe Levi is mistaken to think that reasons may be separated from each other and that all possible combinations should be considered relevant to making a decision for the liberal state. Not only does this postulate hypothetical doctrines that no actual person may endorse, but this may also involve considering doctrines that are dubious or even self-contradictory. For instance, coupling of the probabilistic judgments of a Christian concerning God's existence with the value assessments of an atheist may create an entirely implausible comprehensive doctrine. The Christian's probabilities and evaluations undoubtedly stem from a single comprehensive doctrine concerning God, what counts as evidence for His existence, and His presumed plan for creation. Separating all these and other concerns apart, and then coupling them with the respective views of the atheist suggests a muddled doctrine that no one should take seriously. In the end, reasons and assessment hang together with the connecting thread of a comprehensive doctrine and it does not seem promising to sever this connection for the sake of decision making.

Together, all these reasons suggest to me why the liberal state should not focus on reasons when making a decision. As I understand it, the proper role of deliberation over reasons is to change minds and reveal new options as much as possible. This may go some way for significantly narrowing the field of options that are judged to be optimal according to some permissible value assessment. When deliberation ends and a choice must be made, however, the decision-making process should rely solely on whatever agreements exist between these assessments and not on the reasons underlying these assessments. Even after deliberation, the state may make considerable progress in further narrowing the field of options by eliminating options in accord with considerations like second best versus second worst, weak Pareto, or probabilistic independence, whenever they are applicable. So the liberal state need not dictatorially imposing its judgments in this process, for responsive impartiality by the state is perfectly consistent with the state using agreements concerning the overall value judgments of its constitutive members without having to consider agreements amongst these people on the correct reasons for making the decision.

Of course, when the state must make a decision concerning controversial issues, the set of options regarded as optimal according to some permissible value ordering may be large, even after thoughtful public deliberation and the use of second-, third-, and so on, tiered considerations. At this point, a process like voting is unavoidable to act as the final-tier consideration for the state to ultimately pick amongst these options and choose how to exercise its power. However, in doing so,

the state is not acting on its own comprehensive doctrine. Rather, it is remaining responsively impartial insofar as it is only having its constitutive members vote on options that are regarded as optimal according to some permissible ordering. The will of the people is not revealed, but perhaps at best, an option is selected that will meet with general compliance and support because a significant number of people judge it as the optimal outcome. Furthermore, it should be understood that should public sentiment change, the issue should be open to reconsideration and be reversible to whatever extent possible in the circumstances.

This might seem dissatisfying to ultimately return to voting, especially in light of the concerns from chapter 3, but it should hardly be surprising giving the goal of responsive impartiality. Democracy is a messy process when navigating controversial issues, and making a decision in such circumstances without some public voting process seems positively undemocratic. If the liberal state seeks to remain neutral on these issues, it must give up the aspiration for systematicity in its decisions. Yes, the state may act somewhat arbitrary by appealing to some particular voting method to settle an issue, but it is unavoidable. For making its decisions more systematic and less arbitrary would require the state to begin formulating its own comprehensive doctrine and abandon pretensions of responsive impartiality. The requirement that the results of voting are provisional may lessen this blow, as well as the requirement for certain liberal institutions, for it allows reconsideration of an issue should a social consensus begin to change as well as protecting basic rights and interests of every constitutive member of the state.

To summarize the main thrust of this chapter, dropping the ordering requirement of systematicity and the default rationality thesis removes the blow of Arrow's theorem, but it opens wide the possibilities. I maintain that the liberal state must ultimately choose from the options that are optimal according to at least one permissible value judgment, and that these permissible value judgments need not be the result of interrogating the reasons individuals have behind their individual assessments. Deliberation is valuable in the formation and alternation of these value judgments, but it is these judgments that should be ultimately employed by the liberal state to make its decisions. On controversial social issues, many options may emerge as permissible in this sense, without a consensus on which one of these should be chosen. After all agreements concerning rankings of the options, as well as those concerning additional-tiered considerations have been exhausted, a vote may be necessary in order for the state to pick amongst the permissible options. The dream that a pluralistic society may simply employ a certain mechanical apparatus to discover *the* correct legitimately-binding decision for controversial issues is untenable. The best I can hope for in a messy world is that weak Pareto is adhered to whenever it holds and that an impermissible option, in the sense I give, is not acted upon by the state.

5.5 Concluding Thoughts

The formal results of Arrow's theorem are extremely far-reaching, perhaps further reaching than Arrow himself claims. He states his theorem for the so-called 'non-trivial' cases of more than one person. But, as I argued in chapter 4, even in the 'trivial' case, the theorem still applies: a single person often cannot satisfy the default rationality thesis when he or she has multiple and conflicting value judgments. We have seen theoretical responses in these last two chapters that try to account for coherent decision making in circumstances that the default rationality thesis would abandon as hopelessly arational. The approaches of Sen and Levi flatly reject the norms of rationality presumed by the default rationality thesis while offering their own. It is therefore worth concluding with a reflection on the significance of the many formal results around and about Arrow's theorem.

I am willing to live with state decisions that are not coherent, involve cyclic assessments, or are not completely connected. That would not, of itself, make human society necessarily unjust, nor, as I have maintained, would it necessarily make a sequence of actions by the state irrational. People can live with it, just as they do with their own individual judgments. Our more problematic injustices and irrationalities concern how the alternatives we must choose among are determined and framed; the incongruity between announced social goals and concrete social actions; the commitments we make to act to achieve mutually contradictory goals; how we gather and assess (or ignore and suppress) evidence about the effects of our policies. Formal work in social choice theory asks for clarification concerning notions of rationality and judgment pooling, but are not much help for these more serious problems of social injustice and irrationality. For these, however, there is plenty of room for deliberation.

Bibliography

- Ackerman, Bruce. *Social Justice in the Liberal State* (New Haven, CT: Yale University Press, 1980).
- . ‘Neutralities’, in *Liberalism and the Good* (London: Routledge, 1990): 29–43.
- . *We the People: Foundations* (Cambridge, MA: Belknap Press of Harvard University Press, 1991).
- Ainslie, George. *Picoeconomics: The Strategic Interaction of Successive Motivational States Within the Person* (Cambridge: Cambridge University Press, 1992).
- . *Breakdown of Will* (Cambridge: Cambridge University Press, 2001).
- Aizerman, Mark A. ‘New Problems in the General Choice Theory: Review of a Research Trend’. *Social Choice and Welfare* 2, 4 (December 1985): 235–282.
- Aldred, Jonathan. ‘Social Choice Theory and Deliberative Democracy: A Comment’. *British Journal of Political Science* 34, 4 (October 2004): 747–752.
- Alexander, Sidney S. ‘Social Evaluation through Notional Choice’. *Quarterly Journal of Economics* 88, 4 (1974): 597–624.
- Amadae, Sonja M. *Rationalizing Capitalist Democracy: The Cold War Origins of Rational Choice Liberalism* (Chicago: University of Chicago Press, 2003).
- Anand, Paul. ‘Are the Preference Axioms Really Rational?’ *Theory and Decision* 23, 2 (September 1987): 189–214.
- Anderson, Elizabeth S. *Value in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1993).
- . ‘Practical Reason and Incommensurable Goods’, in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 90–109.
- . ‘What Is the Point of Equality?’ *Ethics* 109, 2 (January 1999): 287–337.

- Anderson, Elizabeth S. 'The Epistemology of Democracy'. *Episteme* 3, 1-2 (June 2006): 8-22.
- Anscombe, F.J. and Robert J. Aumann. 'A Definition of Subjective Probability'. *Annals of Mathematical Statistics* 34, 1 (March 1963): 199-205.
- Aristotle. *The Complete Works of Aristotle: The Revised Oxford Translation*, ed. by Jonathan Barnes (Princeton: Princeton University Press, 1984).
- Arneson, Richard J. 'Equality and Equal Opportunity for Welfare'. *Philosophical Studies* 56, 1 (May 1989): 77-93.
- . 'Defending the Purely Instrumental Account of Democratic Legitimacy'. *Journal of Political Philosophy* 11, 1 (March 2003): 122-132.
- Arrow, Kenneth J. *Social Choice and Individual Values*, 2nd ed. (New York: Wiley, 1963).
- . 'Extended Sympathy and the Possibility of Social Choice'. *American Economic Review* 67, 1 (February 1977): 219-225.
- Arrow, Kenneth J. and Herve Raynaud. *Social Choice and Multicriterion Decision-Making* (Cambridge, MA: MIT Press, 1986).
- Atkinson, Anthony B. 'On the Measurement of Inequality'. *Journal of Economic Theory* 2, 3 (September 1970): 244-263.
- Aumann, Robert J. 'Utility Theory without the Completeness Axiom'. *Econometrica* 30, 3 (July 1962): 445-462.
- . 'Utility Theory without the Completeness Axiom: A Correction'. *Econometrica* 32, 1-2 (January 1964): 210-212.
- Banerjee, Asis and Prasanta K. Pattanaik. 'A Note on a Property of Maximal Sets and Choice in the Absence of Universal Comparability'. *Economics Letters* 51, 2 (May 1996): 191-195.
- Barry, Brian. 'The Public Interest'. *Proceedings of the Aristotelian Society* 38 (1964): 9-14.
- . 'How Not to Defend Liberal Institutions', in *Liberalism and the Good*, ed. by R. Bruce Douglass, Gerald R. Mara and Henry S. Richardson (London: Routledge, 1990): 44-58.
- . *A Treatise on Social Justice*, vol. 2: *Justice as Impartiality* (Oxford: Oxford University Press, 1995).
- Benhabib, Seyla. 'Toward a Deliberative Model of Democratic Legitimacy', in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. by Seyla Benhabib (Princeton: Princeton University Press, 1996): 67-94.
- Bentham, Jeremy. *The Works of Jeremy Bentham*, ed. by John Bowring (London: Simpkin, Marshall, and Company, 1838-1843).
- . *Jeremy Bentham's Economic Writings*, ed. by Werner Stark (London: George Allen and Unwin, 1952-1954).

- . *Deontology together with A Table of the Springs of Action and the Article on Utilitarianism*, ed. by Amnon Goldworth (Oxford: Oxford University Press, 1983).
- Binmore, Ken. *Fun and Games: A Text on Game Theory* (Lexington, MA: D. C. Heath, 1992).
- Black, Duncan. *The Theory of Committees and Elections* (Cambridge: Cambridge University Press, 1958).
- Bohman, James. 'Complexity, Pluralism, and the Constitutional State: On Habermas's *Faktizität und Geltung*'. *Law and Society Review* 28, 4 (1994): 897–930.
- Borda, Jean-Charles de. 'Mémoire sur les Élections au Scrutin', in *Histoire de l'Académie Royale des Sciences* (Paris 1781).
- Bordes, Georges and Nicolaus Tideman. 'Independence of Irrelevant Alternatives in the Theory of Voting'. *Theory and Decision* 39, 2 (March 1991): 163–186.
- Borland, Philip J. 'Majority Systems and the Condorcet Jury Theorem'. *Statistician* 38, 3 (1989): 181–189.
- Broome, John. *Weighing Goods: Equality, Uncertainty and Time* (Oxford: Blackwell, 1991).
- . *Weighing Lives* (Oxford: Oxford University Press, 2004).
- Buchanan, Allen. *Justice, Legitimacy, and Self-Determination* (Oxford: Oxford University Press, 2004).
- Buchanan, James M. 'The Pure Theory of Government Finance: A Suggested Approach'. *Journal of Political Economy* 57, 6 (December 1949): 496–505.
- . 'Individual Choice in Voting and the Market'. *Journal of Political Economy* 62, 4 (August 1954): 334–343.
- . 'Social Choice, Democracy, and Free Markets'. *Journal of Political Economy* 62, 2 (April 1954): 114–123.
- Caplan, Bryan. *The Myth of the Rational Voter: Why Democracies Choose Bad Politics* (Princeton: Princeton University Press, 2007).
- Chang, Ruth. 'Introduction', in *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997): 1–34.
- . 'Putting Together Morality and Well-Being', in *Practical Conflicts: New Philosophical Essays*, ed. by Peter Baumann and Monika Betzler (Cambridge: Cambridge University Press, 2004): 118–158.
- Christiano, Thomas. 'Freedom, Consensus and Equality in Collective Decision Making'. *Ethics* 101, 1 (October 1990): 151–181.
- . 'Social Choice and Democracy', in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 173–195.
- . *Rule of the Many: Fundamental Issues in Democratic Theory* (Boulder, CO: Westview Press, 1996).

- Christiano, Thomas. 'The Significance of Public Deliberation', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 243–277.
- . 'Estlund on Democratic Authority'. *Journal of Political Philosophy* 17, 2 (June 2009): 228–240.
- Cohen, Gerald A. 'On the Currency of Egalitarian Justice'. *Ethics* 99, 4 (July 1989): 906–944.
- Cohen, Joshua. 'An Epistemic Conception of Democracy'. *Ethics* 97, 1 (October 1986): 26–38.
- . 'Reflections on Rousseau: Autonomy and Democracy'. *Philosophy and Public Affairs* 15, 3 (Summer 1986): 275–297.
- . 'Deliberation and Democratic Legitimacy', in *The Good Polity: Normative Analysis of the State*, ed. by Alan Hamlin and Philip Pettit (Oxford: Basil Blackwell, 1989): 17–34.
- . 'Freedom of Expression'. *Philosophy and Public Affairs* 22, 3 (Summer 1993): 207–263.
- . 'Moral Pluralism and Political Consensus', in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 270–291.
- . 'Procedure and Substance in Deliberative Democracy', in *Democracy and Difference: Contesting the Boundaries of the Political*, ed. by Seyla Benhabib (Princeton: Princeton University Press, 1996): 95–119.
- . 'Democracy and Liberty', in *Deliberative Democracy*, ed. by Jon Elster (Cambridge, MA: MIT Press, 1998): 185–231.
- . 'Reflections on Habermas on Democracy'. *Ratio Juris* 12, 4 (December 1999): 385–416.
- Coleman, Jules and John Ferejohn. 'Democracy and Social Choice'. *Ethics* 97, 1 (October 1986): 6–25.
- Condorcet, Marquis de. *Essai sur l'Application de l'Analyse à la Probabilité des Décisions Rendues à la Pluralité des Voix* (Paris: L'Imprimerie Royale, 1785).
- . 'From *An Essay on the Application of Analysis to the Probability of Decisions Rendered by a Plurality of Votes*, 1785', in *Classics of Social Choice*, ed. and trans. by Iain S. McLean and Arnold B. Urken (Ann Arbor: University of Michigan Press, 1995): 91–111.
- Copp, David. 'The Idea of a Legitimate State'. *Philosophy and Public Affairs* 28, 1 (Winter 1999): 3–45.
- Crisp, Roger. 'Equality, Priority, and Compassion'. *Ethics* 113, 4 (July 2003): 745–763.
- Dahl, Robert A. *A Preface to Democratic Theory* (Chicago: University of Chicago Press, 1956).

- . 'Procedural Democracy', in *Philosophy, Politics, and Society: Fifth Series*, ed. by Peter Laslett and James S. Fishkin (New Haven, CT: Yale University Press, 1979): 97–133.
- . *Democracy and Its Critics* (New Haven, CT: Yale University Press, 1989).
- Daniels, Norman. 'Rationing Fairly: Programmatic Considerations'. *Bioethics* 7, 2–3 (1993): 224–233.
- . 'Accountability for Reasonableness: Establishing a Fair Process for Priority Setting is Easier than Agreeing on Principles'. *British Medical Journal* 321, 7272 (November 2000): 1300–1301.
- Daniels, Norman and James Sabin. 'Limits to Health Care: Fair Procedures, Democratic Deliberation, and the Legitimacy Problem for Insurers'. *Philosophy and Public Affairs* 26, 4 (Autumn 1997): 303–350.
- . 'Last Chance Therapies and Managed Care: Pluralism, Fair Procedures, and Legitimacy'. *Hastings Center Report* 28, 2 (March 1998): 27–41.
- . *Setting Limits Fairly: Can We Learn to Share Medical Resources?* (Oxford: Oxford University Press, 2001).
- D'Aspremont, Claude and Louis Gevers. 'Equity and the Informational Basis of Collective Choice'. *Review of Economic Studies* 44, 2 (June 1977): 199–209.
- Davidson, Donald. 'How is Weakness of the Will Possible?', in *Moral Concepts*, ed. by Joel Feinberg (Oxford: Oxford University Press, 1969): 93–113.
- . 'Paradoxes of Irrationality', in *Philosophical Essays on Freud*, ed. by Richard Wollheim and James Hopkins (Cambridge: Cambridge University Press, 1982): 289–305.
- . 'Judging Interpersonal Interests', in *Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 195–211.
- Davidson, Donald, J. C. C. McKinsey and Patrick Suppes. 'Outlines of a Formal Theory of Value, I'. *Philosophy of Science* 22, 2 (April 1955): 140–160.
- Deschamps, Robert and Louis Gevers. 'Leximin and Utilitarian Rules: A Joint Characterization'. *Journal of Economic Theory* 17, 2 (April 1978): 143–163.
- Dewey, John and James H. Tufts. *Ethics*, revised edition (New York: Henry Holt and Company, 1932).
- Dodgson, Charles L. *Facts, Figures, and Fancies, Relating to the Elections to the Hebdomal Council, the Offer of the Clarendon Trustees, and the Proposal to Convert the Parks into Cricket Grounds* (Oxford: Parker, 1874).
- . *The Principles of Parliamentary Representation* (London: Harrison and Sons, 1884).
- Dryzek, John S. *Deliberative Democracy and Beyond: Liberals, Critics, Contestations* (Oxford: Oxford University Press, 2000).

- Dryzek, John S. and Christian List. 'Social Choice Theory and Deliberative Democracy: A Reconciliation'. *British Journal of Political Science* 33, 1 (January 2003): 1–28.
- . 'Social Choice Theory and Deliberative Democracy: A Response to Aldred'. *British Journal of Political Science* 34, 4 (October 2004): 752–758.
- Dummett, Michael. *Voting Procedures* (Oxford: Clarendon Press, 1984).
- Dworkin, Ronald. 'The Original Position'. *University of Chicago Law Review* 40, 3 (Spring 1973): 500–533.
- . 'Did Mill Go Too Far?' *New York Review of Books* 21, 17 (31st October 1974).
- . 'Liberalism', in *Public and Private Morality*, ed. by Stuart Hampshire (Cambridge: Cambridge University Press, 1978): 113–143.
- . *Taking Rights Seriously* (Cambridge, MA: Harvard University Press, 1978).
- . 'Is There a Right to Pornography?' *Oxford Journal of Legal Studies* 1, 2 (Summer 1981): 177–212.
- . 'What is Equality? – Part 2: Equality of Resources'. *Philosophy and Public Affairs* 10, 4 (Autumn 1981): 283–345.
- . 'To Each His Own'. *New York Review of Books* 30, 6 (14th April 1983): 4–6.
- . 'Why Liberals Should Believe in Equality'. *New York Review of Books* 30, 1 (3rd February 1983): 32–34.
- . 'What is Equality? – Part 3: The Place of Liberty'. *Iowa Law Review* 73, 1 (October 1987): 1–54.
- . 'Foundations of Liberal Equality', in *The Tanner Lectures on Human Values*, ed. by Grethe B. Peterson, vol. 11 (Salt Lake City: University of Utah Press, 1990): 1–119.
- . 'Three Questions for America'. *New York Review of Books* 53, 14 (21st September 2006): 24–30.
- Edmundson, William A. 'Legitimate Authority without Political Obligation'. *Law and Philosophy* 17, 1 (January 1998): 43–60.
- Elster, Jon. 'The Case for Methodological Individualism'. *Theory and Society* 11, 4 (July 1982): 453–482.
- . 'The Market and the Forum: Three Varieties of Political Theory', in *The Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 103–132.
- Ely, John Hart. *Democracy and Distrust* (Cambridge, MA: Harvard University Press, 1980).
- Estlund, David M. 'Democracy Without Preference'. *Philosophical Review* 99, 3 (July 1990): 397–423.

- . 'Making Truth Safe for Democracy', in *The Idea of Democracy*, ed. by David Copp, Jean Hampton and John E. Roemer (Cambridge: Cambridge University Press, 1993): 72–100.
- . 'Beyond Fairness and Deliberation: The Epistemic Dimension of Democratic Authority', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 173–204.
- . 'The Insularity of the Reasonable: Why Political Liberalism Must Admit the Truth'. *Ethics* 108, 2 (January 1998): 252–275.
- . 'The Democracy/Contractualism Analogy'. *Philosophy and Public Affairs* 31, 4 (Autumn 2003): 387–412.
- . *Democratic Authority* (Princeton: Princeton University Press, 2008).
- Estlund, David M., Jeremy Waldron, Bernard Grofman and Scott L. Feld. 'Democratic Theory and the Public Interest: Condorcet and Rousseau Revisited'. *American Political Science Review* 83, 4 (December 1989): 1317–1340.
- Fishburn, Peter C. *The Theory of Social Choice* (Princeton: Princeton University Press, 1973).
- Fisher, Roger, William Ury and Bruce Patton. *Gett to Yes: Negotiating Agreement with Giving In*, 2nd ed. (New York: Penguin Books, 1991).
- Fishkin, James S. 'Towards a New Social Contract'. *Noûs* 24, 2 (April 1990): 217–226.
- . *Democracy and Deliberation: New Directions for Democratic Reform* (New Haven, CT: Yale University Press, 1991).
- Frankfurt, Harry. 'Equality as a Moral Ideal'. *Ethics* 98, 1 (October 1987): 21–43.
- Freedman, Benjamin. 'Equipoise and the Ethics of Clinical Research'. *New England Journal of Medicine* 317, 3 (16th July 1987): 141–145.
- Freeman, Samuel. 'Deliberative Democracy: A Sympathetic Comment'. *Philosophy and Public Affairs* 29, 4 (Autumn 2000): 371–418.
- Friedman, Milton. *Capitalism and Freedom* (Chicago: University of Chicago Press, 1962).
- Friedman, Milton and Rose D. Friedman. *Free to Choose: A Personal Statement* (Orlando, FL: Harcourt, 1980).
- Gaertner, Wulf. *A Primer in Social Choice Theory* (Oxford: Oxford University Press, 2006).
- Gibbard, Allan. 'Manipulation of Voting Schemes: A General Result'. *Econometrica* 41, 4 (July 1973): 587–601.
- . 'Interpersonal Comparisons: Preference, Good, and the Intrinsic Reward of a Life', in *Foundations of Social Choice Theory*, ed. by Jon Elster and Aanund Hylland (Cambridge: Cambridge University Press, 1986): 165–193.

- Gilboa, Itzhak and David Schmeidler. 'Maxmin Expected Utility with Non-Unique Prior'. *Journal of Mathematical Economics* 18, 2 (April 1989): 141–153.
- Glymour, Clark. 'Freud's Androids', in *The Cambridge Companion to Freud*, ed. by Jerome Neu (Cambridge: Cambridge University Press, 1991): 44–85.
- Glymour, Clark, David Danks, Bruce Glymour, Frederick Eberhardt, Joseph Ramsey, Richard Scheines, Peter Spirtes, Choh Man Teng and Jiji Zhang. 'Actual Causation: A Stone Soup Essay'. *Synthese* (forthcoming).
- Goodin, Robert E. 'Democratic Deliberation Within'. *Philosophy and Public Affairs* 29, 1 (Winter 2000): 79–107.
- . *Reflective Democracy* (Oxford: Oxford University Press, 2003).
- Goodin, Robert E. and Andrew Reeve. 'Do Neutral Institutions Add Up to a Neutral State?', in *Liberal Neutrality*, ed. by Robert E. Goodin and Andrew Reeve (London: Routledge, 1989): 193–210.
- Goodman, Jay. 'Existence of Compromises in Simple Group Decisions', PhD thesis (Pittsburgh, PA: Department of Statistics, Carnegie Mellon University, 1988).
- Greenawalt, Kent. *Conflicts of Law and Morality* (Oxford: Oxford University Press, 1987).
- Griffin, James. *Well-Being: Its Meaning, Measurement and Moral Importance* (Oxford: Clarendon Press, 1986).
- . 'Incommensurability: What's the Problem?', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 35–51.
- Grofman, Bernard and Scott L. Feld. 'Rousseau's General Will: A Condorcetian Perspective'. *American Political Science Review* 82, 2 (June 1988): 567–576.
- Grofman, Bernard and Guillermo Owen, eds. *Information Pooling and Group Decision Making* (Westport, CT: JAI Press, 1986).
- Grofman, Bernard, Guillermo Owen and Scott L. Feld. 'Thirteen Theorems in Search of Truth'. *Theory and Decision* 15, 3 (September 1983): 261–278.
- Gutmann, Amy and Dennis Thompson. *Democracy and Disagreement* (Cambridge, MA: Belknap Press of Harvard University Press, 1996).
- . 'What Deliberative Democracy Means', in *Why Deliberative Democracy?* (Princeton: Princeton University Press, 2004): 1–63.
- Habermas, Jürgen. *Legitimation Crisis*, trans. by Thomas McCarthy (Boston: Beacon Press, 1975).
- . 'Discourse Ethics: Notes on a Program of Philosophical Justification', in *Moral Consciousness and Communicative Action*, trans. by Christian Lenhardt and Shierry Weber Nicholsen (Cambridge, MA: MIT Press, 1990): 43–115.

- . 'Remarks on Discourse Ethics', in *Justification and Application: Remarks on Discourse Ethics*, trans. by Ciaran P. Cronin (Cambridge, MA: MIT Press, 1993): 19–111.
- . 'Reconciliation through the Public Use of Reason: Remarks on John Rawls's Political Liberalism'. *Journal of Philosophy* 92, 3 (March 1995): 109–131.
- . *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*, trans. by William Rehg (Cambridge, MA: MIT Press, 1996).
- . *The Inclusion of the Other*, ed. by Ciaran P. Cronin and Pablo de Greiff, trans. by Ciaran P. Cronin (Cambridge, MA: MIT Press, 1998).
- Hägele, Günter and Friedrich Pukelsheim. 'Llull's Writings on Electoral Systems'. *Studia Lulliana* 41, 97 (2001): 3–38.
- Hampshire, Stuart. *Innocence and Experience* (Cambridge, MA: Harvard University Press, 1989).
- . 'Liberalism: The New Twist'. *New York Review of Books* 40, 14 (12th August 1993): 43–47.
- Hampton, Jean. 'Contracts and Choices: Does Rawls Have a Social Contract Theory?' *Journal of Philosophy* 77, 6 (June 1980): 315–338.
- Harrison, Ross. *Bentham* (London: Routledge and Kegan Paul, 1983).
- . *Democracy* (London: Routledge, 1993).
- Harsanyi, John C. 'Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking'. *Journal of Political Economy* 61, 5 (October 1953): 434–435.
- . 'Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility'. *Journal of Political Economy* 63, 4 (August 1955): 309–321.
- . 'Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory'. *American Political Science Review* 69, 2 (June 1975): 594–606.
- . 'Bayesian Decision Theory and Utilitarian Ethics'. *American Economic Review* 68, 2 (May 1978): 223–228.
- Hastie, Reid and Robyn M. Dawes. *Rational Choice in an Uncertain World: The Psychology of Judgment and Decision Making* (Thousand Oaks, CA: Sage, 2001).
- Hayek, Friedrich A. 'Scientism and the Study of Society: Part I'. *Economica* 9, 35 (August 1942): 267–291.
- . 'Scientism and the Study of Society: Part II'. *Economica* 10, 37 (February 1943): 34–63.
- . 'Scientism and the Study of Society: Part III'. *Economica* 11, 41 (February 1944): 27–39.
- . *The Constitution of Liberty* (Chicago: University of Chicago Press, 1960).

- Herzberger, Hans G. 'Ordinal Preference and Rational Choice'. *Econometrica* 41, 2 (March 1973): 187–237.
- Hicks, John R. 'The Foundations of Welfare Economics'. *Economic Journal* 49, 196 (December 1939): 696–712.
- Hitchcock, Christopher. 'Of Humean Bondage'. *British Journal for the Philosophy of Science* 54, 1 (2003): 1–25.
- Hobbes, Thomas. *The English Works of Thomas Hobbes of Malmesbury*, ed. by Sir William Molesworth (London: John Bohn, 1839–1845).
- Holmes. 'Dissenting Opinion', in *Lockner v. New York*, 198 U.S. 45, 74 (1905).
- Holmes, Stephen. *The Anatomy of Antiliberalism* (Cambridge, MA: Harvard University Press, 1993).
- Hurka, Thomas. 'Why Value Autonomy?' *Social Theory and Practice* 13, 3 (Fall 1987): 361–382.
- Hurley, Susan L. *Natural Reasons: Personality and Polity* (Oxford: Oxford University Press, 1989).
- Kant, Immanuel. *Practical Philosophy*, ed. and trans. by Mary J. Gregor (Cambridge: Cambridge University Press, 1996).
- Keeney, Ralph L. and Howard Raiffa. *Decisions with Multiple Objectives: Preference and Value Tradeoffs* (Cambridge: Cambridge University Press, 1993).
- Knight, Jack and James Johnson. 'Aggregation and Deliberation: On the Possibility of Democratic Legitimacy'. *Political Theory* 22, 2 (May 1994): 277–296.
- Kornhauser, Lewis A. and Lawrence G. Sage. 'The One and the Many: Adjudication in Collegial Courts'. *California Law Review* 81, 1 (January 1993): 1–59.
- Korsgaard, Christine M. 'Comentary on G. A. Cohen and Amartya Sen', in *The Quality of Life*, ed. by Martha C. Nussbaum and Amartya K. Sen (Oxford: Clarendon Press, 1993): 54–61.
- Kreps, David M. *Notes on the Theory of Choice* (Boulder, CO: Westview Press, 1988).
- Kymlicka, Will. 'Liberal Individualism and Liberal Neutrality'. *Ethics* 99, 4 (July 1989): 883–905.
- . *Liberalism, Community, and Culture* (Oxford: Oxford University Press, 1989).
- Ladenson, Robert. 'In Defense of a Hobbesian Conception of Law'. *Philosophy and Public Affairs* 9, 2 (Winter 1980): 134–59.
- Larmore, Charles E. *Patterns of Moral Complexity* (Cambridge: Cambridge University Press, 1987).
- Lear, Jonathan. 'Inside and Outside the Republic', in *Plato's Republic: Critical Essays*, ed. by Richard Kraut (Lanham, MD: Rowman and Littlefield, 1997): 61–94.

- Levi, Isaac. 'Conflict and Social Agency'. *Journal of Philosophy* 79, 5 (May 1982): 231–247.
- . *Hard Choices: Decision Making Under Unresolved Conflict* (Cambridge: Cambridge University Press, 1986).
- . 'Pareto Unanimity and Consensus'. *Journal of Philosophy* 87, 9 (September 1990): 481–492.
- . 'Conflict and Inquiry'. *Ethics* 102, 4 (July 1992): 814–834.
- . 'Value Commitments, Value Conflict, and the Separability of Belief and Value'. *Philosophy of Science* 66, 4 (December 1999): 509–533.
- . 'Amartya Sen'. *Synthese* 140, 1–2 (May 2004): 61–67.
- . 'Inclusive Rationality'. *Journal of Philosophy* 101, 5 (May 2004): 255–276.
- . 'Why Indeterminate Probability is Rational'. *Journal of Applied Logic* 7, 4 (December 2009): 364–376.
- Lewis, David. 'Causation'. *Journal of Philosophy* 70, 17 (October 1973): 556–567.
- Lipsey, R. G. and Kelvin Lancaster. 'The General Theory of Second Best'. *Review of Economic Studies* 24, 1 (1956–1957): 11–32.
- List, Christian and Robert E. Goodin. 'Epistemic Democracy: Generalizing the Condorcet Jury Theorem'. *Journal of Political Philosophy* 9, 3 (September 2001): 277–306.
- List, Christian, Robert C. Luskin, James S. Fishkin and Iain S. McLean. 'Deliberation, Single-Peakedness, and the Possibility of Meaningful Democracy: Evidence from Deliberative Polls' (unpublished manuscript, 2006).
- List, Christian, Iain S. McLean, James S. Fishkin and Robert C. Luskin. 'Deliberation, Preference Structuration, and Cycles: Evidence from Deliberative Polls' (unpublished manuscript, 2001).
- List, Christian and Philip Pettit. 'Aggregating Sets of Judgments: An Impossibility Result'. *Economics and Philosophy* 18, 1 (April 2002): 89–110.
- . 'Aggregating Sets of Judgments: Two Impossibility Results'. *Synthese* 140, 1–2 (May 2004): 207–235.
- Locke, John. *The Works of John Locke, in Nine Volumes*, 12th ed. (London: C. and J. Rivington, et al., 1824).
- London, Alex John. 'Threats to the Common Good: Biochemical Weapons and Human Subjects Research'. *Hastings Center Report* 33, 5 (September 2003): 17–25.
- Luce, R. Duncan and Howard Raiffa. *Games and Decisions: Introduction and Critical Survey* (New York: Wiley, 1957).
- MacIntyre, Alasdair. *After Virtue: A Study in Moral Theory*, 3rd ed. (Notre Dame: University of Notre Dame Press, 2007).
- Mackie, Gerry. *Democracy Defended* (Cambridge: Cambridge University Press, 2003).

- Manin, Bernard. 'On Legitimacy and Political Deliberation'. *Political Theory* 15, 3 (August 1987): 338–368.
- Mansbridge, Jane. 'A Deliberative Theory of Interest Representation', in *The Politics of Interest*, ed. by Mark Petracca (Boulder, CO: Westview Press, 1992): 32–57.
- Margalit, Avishai. 'Ideals and Second Bests', in *Philosophy for Education*, ed. by Seymour Fox (Jerusalem: Van Leer Foundation, 1983): 77–90.
- May, Simon Căbulea. 'Religious Democracy and the Liberal Principle of Legitimacy'. *Philosophy and Public Affairs* 37, 2 (Spring 2009): 136–170.
- McLean, Iain S. and Fiona Hewitt, eds. *Condorcet: Foundations of Social Choice and Political Theory*, trans. by Iain S. McLean and Fiona Hewitt (Brookfield, VT: Edward Elgar, 1994).
- McLean, Iain S. and Arnold Urken, eds. *Classics of Social Choice*, trans. by Iain S. McLean and Arnold Urken (Ann Arbor: University of Michigan Press, 1995).
- Michelman, Frank I. 'The Subject of Liberalism'. *Stanford Law Review* 46, 6 (July 1994): 1807–1833.
- . 'Review of Jürgen Habermas' *Between Facts and Norms*'. *Journal of Philosophy* 93, 6 (June 1996): 307–315.
- . 'How Can the People Ever Make the Laws? A Critique of Deliberative Democracy', in *Deliberative Democracy: Essays on Reason and Politics*, ed. by James Bohman and William Rehg (Cambridge, MA: MIT Press, 1997): 145–171.
- Mill, John Stuart. *Collected Works of John Stuart Mill*, ed. by J. M. Robson (Toronto: University of Toronto Press, 1963–1991).
- Miller, David. 'Deliberative Democracy and Social Choice'. *Political Studies* 40, special issue (August 1992): 54–67.
- Mongin, Philippe. 'Consistent Bayesian Aggregation'. *Journal of Economic Theory* 66, 2 (August 1995): 313–351.
- Morris, Christopher W. 'Natural Rights and Political Legitimacy'. *Social Philosophy and Policy* 22, 1 (January 2005): 314–329.
- Mueller, Dennis C. *Public Choice III* (Cambridge: Cambridge University Press, 2003).
- Nagel, Thomas. 'Libertarianism without Foundations'. *Yale Law Journal* 85, 1 (November 1975): 136–149.
- . 'Moral Conflict and Political Legitimacy'. *Philosophy and Public Affairs* 16, 3 (Summer 1987): 215–240.
- . *Equality and Partiality* (Oxford: Oxford University Press, 1991).
- Nelson, William. *On Justifying Democracy* (London: Routledge and Kegan Paul, 1980).

- . 'The Very Idea of Pure Procedural Justice'. *Ethics* 90, 4 (July 1980): 502–511.
- Neumann, John von and Oskar Morgenstern. *Theory of Games and Economic Behavior*, 3rd ed. (Princeton: Princeton University Press, 1953).
- Niemi, Richard G. 'Majority Decision-Making with Partial Unidimensionality'. *American Political Science Review* 63, 2 (June 1969): 488–497.
- Nozick, Robert. 'Moral Complications and Moral Structures'. *Natural Law Forum* 13 (1968): 1–50.
- . *Anarchy, State, and Utopia* (New York: Basic Books, 1974).
- . *Philosophical Explanations* (Cambridge, MA: Belknap Press of Harvard University Press, 1981).
- . *The Normative Theory of Individual Choice* (New York: Garland, 1990).
- Nussbaum, Martha C. *Sex and Social Justice* (Oxford: Oxford University Press, 1999).
- . 'Women and Equality: The Capabilities Approach'. *International Labour Review* 138, 3 (September 1999): 227–245.
- . *Women and Human Development* (Cambridge: Cambridge University Press, 2000).
- . 'Forward', in *Libertarianism: For and Against*, ed. by Craig Duncan and Tibor R. Machan (Lanham, MD: Rowman and Littlefield, 2005): ix–xiv.
- Parfit, Derek. 'Equality or Priority', in *The Ideal of Equality*, ed. by Matthew Clayton and Andrew Williams (London: Macmillan, 2000): 81–125.
- Pearce, David G. 'Rationalizable Strategic Behavior and the Problem of Perfection'. *Econometrica* 52, 4 (July 1984): 1029–1050.
- Peterson, Martin. *An Introduction to Decision Theory* (Cambridge: Cambridge University Press, 2009).
- Plato. *Complete Works*, ed. by John M. Cooper and D. S. Hutchinson (Indianapolis, IN: Hackett, 1997).
- Plott, Charles R. 'Path Independence, Rationality, and Social Choice'. *Econometrica* 41, 6 (November 1973): 1075–1091.
- . 'Axiomatic Social Choice Theory: An Overview and Interpretation'. *American Journal of Political Science* 20, 3 (August 1976): 511–596.
- Popper, Karl. 'The Poverty of Historicism, I'. *Economica* 11, 42 (May 1944): 86–103.
- . 'The Poverty of Historicism, II'. *Economica* 11, 43 (August 1944): 119–137.
- . 'The Poverty of Historicism, III'. *Economica* 12, 46 (May 1945): 69–89.
- Putnam, Hilary. 'Über die Rationalität von Präferenzen'. *Allgemeine Zeitschrift für Philosophie* 21, 3 (1996): 209–228.

- Putnam, Hilary. *The Collapse of the Fact/Value Dichotomy and Other Essays* (Cambridge, MA: Harvard University Press, 2002).
- Raiffa, Howard. *Decision Analysis: Introductory Lectures on Choice under Uncertainty* (Reading, MA: Addison-Wesley, 1968).
- Rawls, John. *A Theory of Justice* (Cambridge, MA: Belknap Press of Harvard University Press, 1971).
- . 'Some Reasons for the Maximin Criterion'. *American Economic Review* 64, 2 (May 1974): 141–146.
- . 'The Basic Structure as Subject'. *American Philosophical Quarterly* 14, 2 (April 1977): 159–165.
- . 'Kantian Constructivism in Moral Theory'. *Journal of Philosophy* 77, 9 (September 1980): 515–572.
- . 'The Basic Liberties and Their Priority', in *The Tanner Lectures on Human Values*, ed. by Sterling M. McMurrin, vol. 3 (Salt Lake City: University of Utah Press, 1982): 1–87.
- . 'Social Utility and the Primary Goods', in *Utilitarianism and Beyond*, ed. by Amartya K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982): 159–186.
- . 'Justice as Fairness: Political not Metaphysical'. *Philosophy and Public Affairs* 14, 3 (Summer 1985): 223–251.
- . 'The Idea of a Overlapping Consensus'. *Oxford Journal of Legal Studies* 7, 1 (Spring 1987): 1–25.
- . 'The Priority of Right and Ideas of the Good'. *Philosophy and Public Affairs* 17, 4 (Autumn 1988): 251–276.
- . 'The Domain of the Political and Overlapping Consensus'. *New York University Law Review* 64, 2 (May 1989): 233–255.
- . *Political Liberalism* (New York: Columbia University Press, 1993).
- . 'Introduction to the Paperback Edition', in *Political Liberalism* (New York: Columbia University Press, 1996).
- . 'The Idea of Public Reason Revisited'. *University of Chicago Law Review* 64, 3 (Summer 1997): 765–807.
- . *The Law of Peoples* (Cambridge, MA: Harvard University Press, 1999).
- . *Justice as Fairness: A Restatement*, ed. by Erin Kelly (Cambridge, MA: Harvard University Press, 2001).
- . *Lectures on the History of Political Philosophy*, ed. by Samuel Freeman (Cambridge, MA: Belknap Press of Harvard University Press, 2007).
- Ray, Paramesh. 'Independence of Irrelevant Alternatives'. *Econometrica* 41, 5 (September 1973): 987–991.
- Raz, Joseph. 'Professor Dworkin's Theory of Rights'. *Political Studies* 26, 1 (March 1978): 123–137.

- . *The Morality of Freedom* (Oxford: Oxford University Press, 1986).
- . 'Incommensurability and Agency', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 110–128.
- . *Engaging Reason: On the Theory of Value and Action* (Oxford: Oxford University Press, 1999).
- Regan, Donald. 'Value, Comparability, and Choice', in *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997): 129–150.
- Reidy, David. 'Reciprocity and Reasonable Disagreement: From Liberal to Democratic Legitimacy'. *Philosophical Studies* 132, 2 (January 2007): 243–291.
- Riker, William H. *Democracy in the United States*, 2nd ed. (London: Macmillan, 1964).
- . *Liberalism Against Populism: A Confrontation Between the Theory of Democracy and the Theory of Social Choice* (San Francisco: W. H. Freeman, 1982).
- Robbins, Lionel. *An Essay on the Nature and Significance of Economic Science*, 2nd ed. (London: Macmillan, 1935).
- Roberts, Kevin W. S. 'Interpersonal Comparability and Social Choice Theory'. *Review of Economic Studies* 47, 2 (January 1980): 421–439.
- . 'Possibility Theorems with Interpersonally Comparable Welfare Levels'. *Review of Economic Studies* 47, 2 (January 1980): 409–420.
- Rousseau, Jean-Jacques. *The First and Second Discourses*, ed. by Roger D. Masters, trans. by Roger D. Masters and Judith R. Masters (New York: St. Martin's Press, 1964).
- . *On the Social Contract, with Geneva Manuscript and Political Economy*, ed. by Roger D. Masters, trans. by Judith R. Masters (New York: St. Martin's Press, 1978).
- Rovane, Carol. *The Bounds of Agency: An Essay in Revisionary Metaphysics* (Princeton: Princeton University Press, 1998).
- . 'What is an Agent?' *Synthese* 140, 1–2 (May 2004): 181–198.
- Sandel, Michael J. 'Political Liberalism'. *Harvard Law Review* 107, 7 (May 1994): 1765–1794.
- . *Liberalism and the Limits of Justice*, 2nd ed. (Cambridge: Cambridge University Press, 1998).
- Sartorius, Rolf. 'Political Authority and Political Obligation'. *Virginia Law Review* 67, 1 (February 1981): 3–17.
- Sartre, Jean-Paul. 'Existentialism', in *Existentialism and Human Emotions*, trans. by Bernard Frechtman (New York: Philosophical Library, 1957): 9–51.

- Satterthwaite, Mark A. 'Strategy-Proofness and Arrows Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions'. *Journal of Economic Theory* 10, 2 (1975): 187–217.
- Savage, Leonard J. *The Foundations of Statistics*, 2nd ed. (New York: Dover, 1972).
- Scanlon, Thomas M. 'Contractualism and Utilitarianism', in *Utilitarianism and Beyond*, ed. by Amartya K. Sen and Bernard Williams (Cambridge: Cambridge University Press, 1982): 103–128.
- . *What We Owe to Each Other* (Cambridge, MA: Belknap Press of Harvard University Press, 1998).
- Schervish, Mark J., Teddy Seidenfeld, Joseph B. Kadane and Isaac Levi. 'Extensions of Expected Utility Theory and Some Limitations of Pairwise Comparisons', in *Proceedings of the Third International Symposium on Imprecise Probabilities and Their Applications*, ed. by Jean-Marc Bernard, Teddy Seidenfeld and Marco Zaffalon (Carleton Scientific, 2003): 496–510.
- Schumpeter, Joseph A. *Capitalism, Socialism and Democracy*, 3rd ed. (New York: Harper and Row, 1950).
- Schwartz, Thomas. 'Choice Functions, "Rationality" Conditions, and Variations on the Weak Axiom of Revealed Preference'. *Journal of Economic Theory* 13, 3 (December 1976): 414–427.
- Schwartzberg, Melissa. 'Voting the General Will: Rousseau on Decision Rules'. *Political Theory* 36, 3 (June 2008): 403–423.
- Seidenfeld, Teddy. 'Decision Theory without "Independence" or without "Ordering": What is the Difference?' *Economics and Philosophy* 4, 2 (October 1988): 267–290.
- . 'Outline of a Theory of Partially Ordered Preferences'. *Philosophical Topics* 21, 1 (Spring 1993): 173–189.
- Seidenfeld, Teddy, Joseph B. Kadane and Mark J. Schervish. 'On the Shared Preferences of Two Bayesian Decision Makers'. *Journal of Philosophy* 86, 5 (May 1989): 225–244.
- Seidenfeld, Teddy, Mark J. Schervish and Joseph B. Kadane. 'A Representation of Partially Ordered Preferences'. *Annals of Statistics* 23, 6 (December 1995): 2168–2217.
- . 'Coherent Choice Functions Under Uncertainty'. *Synthese* 172, 1 (January 2010): 157–176.
- Sen, Amartya K. 'A Possibility Theorem on Majority Decisions'. *Econometrica* 34, 2 (April 1966): 491–499.
- . 'Quasi-Transitivity, Rational Choice and Collective Decisions'. *Review of Economic Studies* 36, 3 (July 1969): 381–393.
- . *Collective Choice and Social Welfare* (Amsterdam: North-Holland, 1970).

- . 'Interpersonal Aggregation and Partial Comparability'. *Econometrica* 38, 3 (May 1970): 393–409.
- . 'Choice Functions and Revealed Preference'. *Review of Economic Studies* 38, 3 (July 1971): 307–317.
- . 'Interpersonal Comparison and Partial Comparability: A Correction'. *Econometrica* 40, 5 (September 1972): 959.
- . *On Economic Inequality* (Oxford: Clarendon Press, 1973).
- . 'Social Choice Theory: A Re-Examination'. *Econometrica* 45, 1 (January 1977): 53–89.
- . 'Interpersonal Comparisons of Welfare', in *Economics and Human Welfare: Essays in Honor of Tibor Scitovsky*, ed. by Michael J. Boskin (New York: Academic Press, 1979): 183–201.
- . 'Equality of What?', in *Tanner Lectures on Human Values*, ed. by Sterling M. McMurrin, vol. 1 (Salt Lake City: University of Utah Press, 1980): 195–220.
- . *Commodities and Capabilities* (Amsterdam: North-Holland, 1985).
- . 'Rationality and Uncertainty'. *Theory and Decision* 18, 2 (March 1985): 109–127.
- . *On Ethics and Economics* (Oxford: Blackwell, 1987).
- . *Inequality Reexamined* (Oxford: Oxford University Press, 1992).
- . 'Internal Consistency of Choice'. *Econometrica* 61, 3 (May 1993): 495–521.
- . 'Non-Binary Choice and Preference', in *Logic, Methodology and Philosophy of Science IX*, ed. by Dag Prawitz, Brian Skyrms and Dag Westerstaahl (Amsterdam: Elsevier Science, 1994).
- . 'The Formulation of Rational Choice'. *American Economic Review* 84, 2 (May 1994): 385–390.
- . 'Rationality and Social Choice'. *American Economic Review* 85, 1 (March 1995): 1–24.
- . 'Maximization and the Act of Choice'. *Econometrica* 65, 4 (July 1997): 745–779.
- . 'Incompleteness and Reasoned Choice'. *Synthese* 140, 1–2 (May 2004): 43–59.
- Sher, George. *Beyond Neutrality: Perfectionism and Politics* (Cambridge: Cambridge University Press, 1997).
- Sigwick, Henry. *The Methods of Ethics*, 7th ed. (London: Macmillan, 1907).
- Simmons, A. John. *Moral Principles and Political Obligations* (Princeton: Princeton University Press, 1979).
- . 'Liberal Impartiality and Political Legitimacy'. *Political Books* 34, 4 (October 1993): 213–223.

- Simmons, A. John. 'Justification and Legitimacy'. *Ethics* 109, 4 (July 1999): 739–771.
- . 'Political Obligation and Authority', in *The Blackwell Guide to Social and Political Philosophy*, ed. by Robert L. Simon (Oxford: Blackwell, 2002): 17–37.
- . 'The Duty to Obey and Our Natural Moral Duties', in *Is There a Duty to Obey the Law? For and Against*, ed. by Christopher Heath Wellman and A. John Simmons (Cambridge: Cambridge University Press, 2005): 91–196.
- Smith, M. B. E. 'Is There a Prima Facie Obligation to Obey the Law?' *Yale Law Journal* 82, 5 (April 1973): 950–976.
- Stark, Cynthia A. 'Hypothetical Consent and Justification'. *Journal of Philosophy* 97, 6 (June 2000): 313–334.
- Steinhaus, Hugo. 'The Problem of Fair Division'. *Econometrica* 16, 1 (January 1948): 101–104.
- Stocker, Michael. *Plural and Conflicting Values* (Oxford: Clarendon Press, 1990).
- Sunstein, Cass. 'Beyond the Republican Revival'. *Yale Law Journal* 97, 8 (July 1988): 1539–1590.
- . *Democracy and the Problem of Free Speech* (New York: The Free Press, 1993).
- Surowiecki, James. *The Wisdom of Crowds: Why the Many Are Smarter Than the Few and How Collective Wisdom Shapes Business* (New York: Doubleday, 2004).
- Suzumura, Kotaro. 'Remarks on the Theory of Collective Choice'. *Economica* 43, 172 (November 1976): 381–390.
- . *Rational Choice, Collective Decisions, and Social Welfare* (Cambridge: Cambridge University Press, 1983).
- Suzumura, Kotaro and Yongsheng Xu. 'Recoverability of Choice Functions and Binary Relations: Some Duality Results'. *Social Choice and Welfare* 21, 1 (August 2003): 21–37.
- Taylor, Charles. 'What's Wrong with Negative Liberty?', in *The Idea of Freedom: Essays in Honour of Isaiah Berlin*, ed. by Alan Ryan (Oxford: Oxford University Press, 1979): 175–193.
- . 'The Diversity of Goods', in *Utilitarianism and Beyond*, ed. by Bernard Williams and Amartya K. Sen (Cambridge: Cambridge University Press, 1982): 129–144.
- . 'Leading a Life', in *Incommensurability, Incomparability, and Practical Reason*, ed. by Ruth Chang (Cambridge, MA: Harvard University Press, 1997): 170–183.
- Temkin, Larry S. *Inequality* (Oxford: Oxford University Press, 1996).
- . 'Egalitarianism Defended'. *Ethics* 113, 4 (July 2003): 764–782.

- Thomson, Judith Jarvis. 'Some Ruminations on Rights'. *Arizona Law Review* 19, 1 (1977): 45–60.
- Tversky, Amos and Daniel Kahneman. 'Judgement Under Uncertainty: Heuristics and Biases'. *Science* 185, 4157 (1974): 1124–1131.
- Waldron, Jeremy. 'John Rawls and the Social Minimum'. *Journal of Applied Philosophy* 3, 1 (March 1986): 21–33.
- . 'Theoretical Foundations of Liberalism'. *Philosophical Quarterly* 37, 147 (April 1987): 127–150.
- . 'Legislation and Moral Neutrality', in *Liberal Neutrality*, ed. by Robert E. Goodin and Andrew Reeve (London: Routledge, 1989): 61–83.
- Wall, Steven. 'Democracy and Equality'. *Philosophical Quarterly* 57, 228 (July 2007): 416–438.
- Walley, Peter. *Statistical Reasoning with Imprecise Probabilities* (London: Chapman and Hall, 1991).
- Walzer, Michael. *Spheres of Justice: A Defense of Pluralism and Equality* (New York: Basic Books, 1983).
- Williams, Bernard. *Ethics and the Limits of Philosophy* (Cambridge: Cambridge University Press, 1985).
- Wolin, Sheldon S. 'Democracy, Difference, and Re-Cognition'. *Political Theory* 21, 3 (August 1993): 464–483.
- Young, Iris Marion. *Justice and the Politics of Difference* (Princeton: Princeton University Press, 1990).